

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the application of: J. Don Chen

Serial No.: N/A

Filed: Herewith

For: *NOVEL NUCLEAR RECEPTOR
COREPRESSOR MOLECULES AND USES
THEREFOR*

Attorney Docket No.: UMG-030

Group Art Unit: Unassigned

Examiner: Unassigned

Box Sequence
Commissioner for Patents
Washington, D.C. 20231

**TRANSMITTAL LETTER FOR DISKETTE CONTAINING SEQUENCE
LISTING**

Dear Sir:

Enclosed is a diskette which contains a computer readable form of the Sequence Listing for the patent application filed herewith. The Sequence Listing complies with the requirements of 37 C.F.R. § 1.821. The information recorded in computer readable form is identical to the written Sequence Listing appearing on pages 1-59 submitted herewith, as required by 37 C.F.R. § 1.821(f). The computer readable form of the Sequence Listing contained on the enclosed diskette is understood to comply with the requirements of § 1.824(d).

"Express Mail" mailing label number EL 746 157 131 US

Date of Deposit March 27, 2001

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner for Patents, Box Patent Application, Washington, D.C. 20231

Viriato G. Cardoso
Signature

Viriato G. Cardoso

Please Print Name of Person Signing

LAHIVE & COCKFIELD, LLP
Attorneys at Law

By *Peter C. Lauro*
Peter C. Lauro, Esq.
Reg. No. 32,360
28 State Street
Boston, MA 02109
Telephone: 617-227-7400
Facsimile: 617-742-4214

SEQUENCE LISTING

<110> Chen, Don J.

<120> NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
AND USES THEREFOR

<130> UMG-030

<140>

<141>

<150> 60/193,138

<151> 2000-03-29

<160> 6

<170> PatentIn Ver. 2.0

<210> 1

<211> 8686

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (157)..(7677)

<400> 1

gagtcctttga ggacacagcc tcgctggagg cagtttcttg tgccagtgac ggggtggccc 60

gtgagctgat gacgaggact ggcttttaat ccttggtggt gattaagaga aagcttattg 120

gggcctggga gcagctcccc gccgaccccc accacc atg tcg ggc tcc aca cag 174
Met Ser Gly Ser Thr Gln
1 5

cct gtg gca cag acg tgg agg gcc act gag ccc cgc tac ccg ccc cac 222
Pro Val Ala Gln Thr Trp Arg Ala Thr Glu Pro Arg Tyr Pro Pro His
10 15 20

agc ctt tcc tac cca gtg cag atc gcc cgg acg cac acg gac gtc ggg 270
Ser Leu Ser Tyr Pro Val Gln Ile Ala Arg Thr His Thr Asp Val Gly
25 30 35

ctc ctg gag tac cag cac cac tcc cgc gac tat gcc tcc cac ctg tcg 318
Leu Leu Glu Tyr Gln His His Ser Arg Asp Tyr Ala Ser His Leu Ser
40 45 50

ccc ggc tcc atc atc cag ccc cag cgg cgg agg ccc tcc ctg ctg tct 366
Pro Gly Ser Ile Ile Gln Pro Gln Arg Arg Arg Pro Ser Leu Leu Ser
55 60 65 70

gag ttc cag ccc ggg aat gaa cgg tcc cag gag ctc cac ctg cgg cca 414
Glu Phe Gln Pro Gly Asn Glu Arg Ser Gln Glu Leu His Leu Arg Pro
75 80 85

gag tcc cac tca tac ctg ccc gag ctg ggg aag tca gag atg gag ttc 462
Glu Ser His Ser Tyr Leu Pro Glu Leu Gly Lys Ser Glu Met Glu Phe
90 95 100

PatentIn Ver. 2.0

att gaa agc aag cgc cct cgg cta gag ctg ctg cct gac ccc ctg ctg	510
Ile Glu Ser Lys Arg Pro Arg Leu Glu Leu Leu Pro Asp Pro Leu Leu	
105 110 115	
cga ccg tca ccc ctg ctg gcc acg ggc cag cct gcg gga tct gaa gac	558
Arg Pro Ser Pro Leu Leu Ala Thr Gly Gln Pro Ala Gly Ser Glu Asp	
120 125 130	
ctc acc aag gac cgt agc ctg acg ggc aag ctg gaa ccg gtg tct ccc	606
Leu Thr Lys Asp Arg Ser Leu Thr Gly Lys Leu Glu Pro Val Ser Pro	
135 140 145 150	
ccc agc ccc ccg cac act gac cct gag ctg gag ctg gtg ccg cca cgg	654
Pro Ser Pro Pro His Thr Asp Pro Glu Leu Glu Leu Val Pro Pro Arg	
155 160 165	
ctg tcc aag gag gag ctg atc cag aac atg gac cgc gtg gac cga gag	702
Leu Ser Lys Glu Glu Leu Ile Gln Asn Met Asp Arg Val Asp Arg Glu	
170 175 180	
atc acc atg gta gag cag cag atc tct aag ctg aag aag aag cag caa	750
Ile Thr Met Val Glu Gln Gln Ile Ser Lys Leu Lys Lys Lys Gln Gln	
185 190 195	
cag ctg gag gag gag gct gcc aag ccg ccc gag cct gag aag ccc gtg	798
Gln Leu Glu Glu Glu Ala Ala Lys Pro Pro Glu Pro Glu Lys Pro Val	
200 205 210	
tca ccg ccg ccc atc gag tgc aag cac cgc agc ctg gtg cag atc atc	846
Ser Pro Pro Pro Ile Glu Ser Lys His Arg Ser Leu Val Gln Ile Ile	
215 220 225 230	
tac gac gag aac cgg aag aag gct gaa gct gca cat cgg att ctg gaa	894
Tyr Asp Glu Asn Arg Lys Lys Ala Glu Ala Ala His Arg Ile Leu Glu	
235 240 245	
ggc ctg ggg ccc cag gtg gag ctg ccg ctg tac aac cag ccc tcc gac	942
Gly Leu Gly Pro Gln Val Glu Leu Pro Leu Tyr Asn Gln Pro Ser Asp	
250 255 260	
acc ccg cag tat cat gag aac atc aaa ata aac cag gcg atg ccg aag	990
Thr Arg Gln Tyr His Glu Asn Ile Lys Ile Asn Gln Ala Met Arg Lys	
265 270 275	
aag cta atc ttg tac ttc aag agg agg aat cac gct ccg aaa caa tgg	1038
Lys Leu Ile Leu Tyr Phe Lys Arg Arg Asn His Ala Arg Lys Gln Trp	
280 285 290	
gag cag aag ttc tgc cag cgc tat gac cag ctc atg gag gcc tgg gag	1086
Glu Gln Lys Phe Cys Gln Arg Tyr Asp Gln Leu Met Glu Ala Trp Glu	
295 300 305 310	
aag aag gtg gag cgc atc gag aac aac ccc ccg ccg ccg gcc aag gag	1134
Lys Lys Val Glu Arg Ile Glu Asn Asn Pro Arg Arg Arg Ala Lys Glu	
315 320 325	
agc aag gtt cgc gag tac tac gag aag cag ttc cct gag atc cgc aag	1182
Ser Lys Val Arg Glu Tyr Tyr Glu Lys Gln Phe Pro Glu Ile Arg Lys	
330 335 340	
cag cgc gag ctg cag gag cgc atg cag agg gtg gcc cag ccg gcc agt	1230

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gln	Arg	Glu	Leu	Gln	Glu	Arg	Met	Gln	Arg	Val	Gly	Gln	Arg	Gly	Ser	
		345					350					355				
ggg	ctg	tcc	atg	tcg	ccc	gcc	cgc	agc	gag	cac	gag	gtg	tca	gag	atc	1278
Gly	Leu	Ser	Met	Ser	Pro	Ala	Arg	Ser	Glu	His	Glu	Val	Ser	Glu	Ile	
	360					365					370					
atc	gat	ggc	ctc	tca	gag	cag	gag	aac	ctg	gag	aag	cag	atg	cgc	cag	1326
Ile	Asp	Gly	Leu	Ser	Glu	Gln	Glu	Asn	Leu	Glu	Lys	Gln	Met	Arg	Gln	
375					380					385					390	
ctg	gcc	gtg	atc	ccg	ccc	atg	ctg	tac	gac	gct	gac	cag	cag	cgc	atc	1374
Leu	Ala	Val	Ile	Pro	Pro	Met	Leu	Tyr	Asp	Ala	Asp	Gln	Gln	Arg	Ile	
				395					400					405		
aag	ttc	atc	aac	atg	aac	ggg	ctt	atg	gcc	gac	ccc	atg	aag	gtg	tac	1422
Lys	Phe	Ile	Asn	Met	Asn	Gly	Leu	Met	Ala	Asp	Pro	Met	Lys	Val	Tyr	
		410					415						420			
aaa	gac	cgc	cag	gtc	atg	aac	atg	tgg	agt	gag	cag	gag	aag	gag	acc	1470
Lys	Asp	Arg	Gln	Val	Met	Asn	Met	Trp	Ser	Glu	Gln	Glu	Lys	Glu	Thr	
		425					430					435				
ttc	cgg	gag	aag	ttc	atg	cag	cat	ccc	aag	aac	ttt	ggc	ctg	atc	gca	1518
Phe	Arg	Glu	Lys	Phe	Met	Gln	His	Pro	Lys	Asn	Phe	Gly	Leu	Ile	Ala	
	440					445					450					
tca	ttc	ctg	gag	agg	aag	aca	gtg	gct	gag	tgc	gtc	ctc	tat	tac	tac	1566
Ser	Phe	Leu	Glu	Arg	Lys	Thr	Val	Ala	Glu	Cys	Val	Leu	Tyr	Tyr	Tyr	
455					460					465					470	
ctg	act	aag	aag	aat	gag	aac	tat	aag	agc	ctg	gtg	aga	cgg	agc	tat	1614
Leu	Thr	Lys	Lys	Asn	Glu	Asn	Tyr	Lys	Ser	Leu	Val	Arg	Arg	Ser	Tyr	
				475					480					485		
cgg	cgc	cgc	ggc	aag	agc	cag	cag	caa	caa	cag	cag	cag	cag	cag	cag	1662
Arg	Arg	Arg	Gly	Lys	Ser	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	
			490					495					500			
cag	cag	cag	cag	cag	cag	cag	ccc	atg	ccc	cgc	agc	agc	cag	gag	gag	1710
Gln	Gln	Gln	Gln	Gln	Gln	Gln	Pro	Met	Pro	Arg	Ser	Ser	Gln	Glu	Glu	
		505					510					515				
aaa	gat	gag	aag	gag	aag	gaa	aag	gag	gcg	gag	aag	gag	gag	gag	aag	1758
Lys	Asp	Glu	Lys	Glu	Lys	Glu	Lys	Glu	Ala	Glu	Lys	Glu	Glu	Glu	Lys	
	520					525					530					
ccg	gag	gtg	gag	aac	gac	aag	gaa	gac	ctc	ctc	aag	gag	aag	aca	gac	1806
Pro	Glu	Val	Glu	Asn	Asp	Lys	Glu	Asp	Leu	Leu	Lys	Glu	Lys	Thr	Asp	
535					540					545					550	
gac	acc	tca	ggg	gag	gac	aac	gac	gag	aag	gag	gct	gtg	gcc	tcc	aaa	1854
Asp	Thr	Ser	Gly	Glu	Asp	Asn	Asp	Glu	Lys	Glu</						

	585					590					595					
cag Gln	agc Ser	gcc Ala	gag Glu	ctg Leu	gcc Ala	tcc Ser	atg Met	gag Glu	ctg Leu	aat Asn	gag Glu	agt Ser	tct Ser	cgc Arg	tgg Trp	1998
600					605					610						
aca Thr	gaa Glu	gaa Glu	gaa Glu	atg Met	gaa Glu	aca Thr	gcc Ala	aag Lys	aaa Lys	ggt Gly	ctc Leu	ctg Leu	gaa Glu	cac His	ggc Gly	2046
615					620					625						
cgc Arg	aac Asn	tgg Trp	tcg Ser	gcc Ala	atc Ile	gcc Ala	cgg Arg	atg Met	gtg Val	ggc Gly	tcc Ser	aag Lys	act Thr	gtg Val	tcg Ser	2094
635					640					645						
cag Gln	tgt Cys	aag Lys	aac Asn	ttc Phe	tac Tyr	ttc Phe	aac Asn	tac Tyr	aag Lys	aag Lys	agg Arg	cag Gln	aac Asn	ctc Leu	gat Asp	2142
650					655					660						
gag Glu	atc Ile	ttg Leu	cag Gln	cag Gln	cac His	aag Lys	ctg Leu	aag Lys	atg Met	gag Glu	aag Lys	gag Glu	agg Arg	aac Asn	gcg Ala	2190
665					670					675						
cgg Arg	agg Arg	aag Lys	aag Lys	aag Lys	aaa Lys	gcg Ala	ccg Pro	gcg Ala	gcg Ala	gcc Ala	agc Ser	gag Glu	gag Glu	gct Ala	gca Ala	2238
680					685					690						
ttc Phe	ccg Pro	ccc Pro	gtg Val	gtg Val	gag Glu	gat Asp	gag Glu	gag Glu	atg Met	gag Glu	gcg Ala	tcg Ser	ggc Gly	gtg Val	acg Thr	2286
695					700					705						
gga Gly	aat Asn	gag Glu	gag Glu	gag Glu	atg Met	gtg Val	gag Glu	gag Glu	gct Ala	gaa Glu	gcc Ala	act Thr	gtc Val	aac Asn	aac Asn	2334
715					720					725						
agc Ser	tca Ser	gac Asp	acc Thr	gag Glu	agc Ser	atc Ile	ccc Pro	tct Ser	cct Pro	cac His	act Thr	gag Glu	gcc Ala	gcc Ala	aag Lys	2382
730					735					740						
gac Asp	aca Thr	ggg Gly	cag Gln	aat Asn	ggg Gly	ccc Pro	aag Lys	ccc Pro	cca Pro	gcc Ala	acc Thr	ctg Leu	ggc Gly	gcc Ala	gac Asp	2430
745					750					755						
ggg Gly	cca Pro	ccc Pro	cca Pro	ggg Gly	cca Pro	ccc Pro	acc Thr	cca Pro	cca Pro	ccg Pro	gag Glu	gac Asp	atc Ile	ccg Pro	gcc Ala	2478
760					765					770						
ccc Pro	act Thr	gag Glu	tcc Ser	acc Thr	ccg Pro	gcc Ala	tct Ser	gaa Glu	gcc Ala	acc Thr	tta Leu	gcc Ala	cct Pro	acg Thr	ccc Pro	2526
775					780					785						
cca Pro	cca Pro	gca Ala	ccc Pro	cca Pro	ttt Phe	ccc Pro	tct Ser	tca Ser	cct Pro	cct Pro	cct Pro	gtg Val	gtc Val	ccc Pro	aag Lys	2574
795					800					805						
gag Glu	gag Glu	aag Lys	gag Glu	gag Glu	gag Glu	acc Thr	gca Ala	gca Ala	gcg Ala	ccc Pro	cca Pro	gtg Val	gag Glu	gag Glu	ggg Gly	2622
810					815					820						
gag Glu	gag Glu	cag Gln	aag Lys	ccc Pro	ccc Pro	gcg Ala	gct Ala	gag Glu	gag Glu	ctg Leu	gca Ala	gtg Val	gac Asp	aca Thr	ggg Gly	2670
825					830					835						

aag Lys	gcc Ala	gag Glu	gag Glu	ccc Pro	gtc Val	aag Lys	agc Ser	gag Glu	tgc Cys	acg Thr	gag Glu	gaa Glu	gcc Ala	gag Glu	gag Glu	2718
840 845 850																
ggg Gly	ccg Pro	gcc Ala	aag Lys	ggc Gly	aag Lys	gac Asp	gcg Ala	gag Glu	gcc Ala	gct Ala	gag Glu	gcc Ala	acg Thr	gcc Ala	gag Glu	2766
855 860 865 870																
agg Arg	gcg Ala	ctc Leu	aag Lys	gca Ala	gag Glu	aag Lys	aag Lys	gag Glu	ggc Gly	ggg Gly	agc Ser	ggc Gly	agg Arg	gcc Ala	acc Thr	2814
875 880 885																
aca Thr	gcc Ala	aag Lys	agc Ser	tcg Ser	ggc Gly	gcc Ala	ccc Pro	cag Gln	gac Asp	agc Ser	gac Asp	tcc Ser	agt Ser	gcc Ala	acc Thr	2862
890 895 900																
tgc Cys	agt Ser	gca Ala	gac Asp	gag Glu	gtg Val	gat Asp	gag Glu	gcc Ala	gag Glu	ggc Gly	ggc Gly	gac Asp	aag Lys	aac Asn	cgg Arg	2910
905 910 915																
ctg Leu	ctg Leu	tcc Ser	cca Pro	agg Arg	ccc Pro	agc Ser	ctc Leu	ctc Leu	acc Thr	ccg Pro	act Thr	ggc Gly	gac Asp	ccc Pro	cgg Arg	2958
920 925 930																
gcc Ala	aat Asn	gcc Ala	tca Ser	ccc Pro	cag Gln	aag Lys	cca Pro	ctg Leu	gac Asp	ctg Leu	aag Lys	cag Gln	ctg Leu	aag Lys	cag Gln	3006
935 940 945 950																
cga Arg	gcg Ala	gct Ala	gcc Ala	atc Ile	ccc Pro	ccc Pro	atc Ile	cag Gln	gtc Val	acc Thr	aaa Lys	gtc Val	cat His	gag Glu	ccc Pro	3054
955 960 965																
ccc Pro	cgg Arg	gag Glu	gac Asp	gca Ala	gct Ala	ccc Pro	acc Thr	aag Lys	cca Pro	gct Ala	ccc Pro	cca Pro	gcc Ala	cca Pro	cgg Pro	3102
970 975 980																
cca Pro	ccg Pro	caa Gln	aac Asn	ctg Leu	cag Gln	ccg Pro	gag Glu	agc Ser	gac Asp	gcc Ala	cct Pro	cag Gln	cag Gln	cct Pro	ggc Gly	3150
985 990 995																
agc Ser	agc Ser	ccc Pro	cgg Arg	ggc Gly	aag Lys	agc Ser	agg Arg	agc Ser	ccg Pro	gca Ala	ccc Pro	ccc Pro	gcc Ala	gac Asp	aag Lys	3198
1000 1005 1010																
gag Glu	gca Ala	gag Glu	aag Lys	cct Pro	gtg Val	ttc Phe	ttc Phe	cca Pro	gcc Ala	ttc Phe	gca Ala	gcc Ala	gag Glu	gcc Ala	cag Gln	3246
1015 1020 1025 1030																
aag Lys	ctg Leu	cct Pro	ggg Gly	gac Asp	ccc Pro	cct Pro	tgc Cys	tgg Trp	act Thr	tcc Ser	ggc Gly	ctg Leu	ccc Pro	ttc Phe	ccc Pro	3294
1035 1040 1045																
gtg Val	ccc Pro	ccc Pro	cgt Arg	gag Glu	gtg Val	atc Ile	aag Lys	gcc Ala	tcc Ser	ccg Pro	cat His	gcc Ala	ccg Pro	gac Asp	ccc Pro	3342
1050 1055 1060																
tca Ser	gcc Ala	ttc Phe	tcc Ser	tac Tyr	gct Ala	cca Pro	cct Pro	ggt Gly	cac His	cca Pro	ctg Leu	ccc Pro	ctg Leu	ggc Gly	ctc Leu	3390
1065 1070 1075																

cat gac act gcc cgg ccc gtc ctg ccg cgc cca ccc acc atc tcc aac	3438
His Asp Thr Ala Arg Pro Val Leu Pro Arg Pro Pro Thr Ile Ser Asn	
1080 1085 1090	
ccg cct ccc ctc atc tcc tct gcc aag cac ccc agc gtc ctc gag agg	3486
Pro Pro Pro Leu Ile Ser Ser Ala Lys His Pro Ser Val Leu Glu Arg	
1095 1100 1105 1110	
caa ata ggt gcc atc tcc caa gga atg tgc gtc cag ctc cac gtc ccg	3534
Gln Ile Gly Ala Ile Ser Gln Gly Met Ser Val Gln Leu His Val Pro	
1115 1120 1125	
tac tca gag cat gcc aag gcc ccg gtg ggc cct gtc acc atg ggg ctg	3582
Tyr Ser Glu His Ala Lys Ala Pro Val Gly Pro Val Thr Met Gly Leu	
1130 1135 1140	
ccc ctg ccc atg gac ccc aaa aag ctg gca ccc ttc agc gga gtg aag	3630
Pro Leu Pro Met Asp Pro Lys Lys Leu Ala Pro Phe Ser Gly Val Lys	
1145 1150 1155	
cag gag cag ctg tcc cca cgg ggc cag gct ggg cca ccg gag agc ctg	3678
Gln Glu Gln Leu Ser Pro Arg Gly Gln Ala Gly Pro Pro Glu Ser Leu	
1160 1165 1170	
ggg gtg ccc aca gcc cag gag gcg tcc gtg ctg aga ggg aca gct ctg	3726
Gly Val Pro Thr Ala Gln Glu Ala Ser Val Leu Arg Gly Thr Ala Leu	
1175 1180 1185 1190	
ggc tca gtt ccg ggc gga agc atc acc aaa ggc att ccc agc aca cgg	3774
Gly Ser Val Pro Gly Gly Ser Ile Thr Lys Gly Ile Pro Ser Thr Arg	
1195 1200 1205	
gtg ccc tgc gac agc gcc atc aca tac cgc ggc tcc atc acc cac ggc	3822
Val Pro Ser Asp Ser Ala Ile Thr Tyr Arg Gly Ser Ile Thr His Gly	
1210 1215 1220	
acg cca gct gac gtc ctg tac aag ggc acc atc acc agg atc atc ggc	3870
Thr Pro Ala Asp Val Leu Tyr Lys Gly Thr Ile Thr Arg Ile Ile Gly	
1225 1230 1235	
gag gac agc ccg agt cgc ttg gac cgc ggc cgg gag gac agc ctg ccc	3918
Glu Asp Ser Pro Ser Arg Leu Asp Arg Gly Arg Glu Asp Ser Leu Pro	
1240 1245 1250	
aag ggc cac gtc atc tac gaa ggc aag aag ggc cac gtc ttg tcc tat	3966
Lys Gly His Val Ile Tyr Glu Gly Lys Lys Gly His Val Leu Ser Tyr	
1255 1260 1265 1270	
gag ggt ggc atg tct gtg acc cag tgc tcc aag gag gac ggc aga agc	4014
Glu Gly Gly Met Ser Val Thr Gln Cys Ser Lys Glu Asp Gly Arg Ser	
1275 1280 1285	
agc tca gga ccc ccc cat gag acg gcc gcc ccc aag cgc acc tat gac	4062
Ser Ser Gly Pro Pro His Glu Thr Ala Ala Pro Lys Arg Thr Tyr Asp	
1290 1295 1300	
atg atg gag ggc cgc gtg ggc aga gcc atc tcc tca gcc agc atc gaa	4110
Met Met Glu Gly Arg Val Gly Arg Ala Ile Ser Ser Ala Ser Ile Glu	
1305 1310 1315	
ggt ctc atg ggc cgt gcc atc ccg ccg gag cga cac agc ccc cac cac	4158

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted November 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gly Leu Met Gly Arg Ala Ile Pro Pro Glu Arg His Ser Pro His His			
1320	1325	1330	
ctc aaa gag cag cac cac atc cgc ggg tcc atc aca caa ggg atc cct			4206
Leu Lys Glu Gln His His Ile Arg Gly Ser Ile Thr Gln Gly Ile Pro			
1335	1340	1345	1350
cgg tcc tac gtg gag gca cag gag gac tac ctg cgt cgg gag gcc aag			4254
Arg Ser Tyr Val Glu Ala Gln Glu Asp Tyr Leu Arg Arg Glu Ala Lys			
1355	1360	1365	
ctc cta aag cgg gag ggc acg cct ccg ccc cca ccg ccc tca cgg gac			4302
Leu Leu Lys Arg Glu Gly Thr Pro Pro Pro Pro Pro Pro Ser Arg Asp			
1370	1375	1380	
ctg acc gag gcc tac aag acg cag gcc ctg ggc ccc ctg aag ctg aag			4350
Leu Thr Glu Ala Tyr Lys Thr Gln Ala Leu Gly Pro Leu Lys Leu Lys			
1385	1390	1395	
ccg gcc cat gag ggc ctg gtg gcc acg gtg aag gag gcg ggc cgc tcc			4398
Pro Ala His Glu Gly Leu Val Ala Thr Val Lys Glu Ala Gly Arg Ser			
1400	1405	1410	
atc cat gag atc ccg cgc gag gag ctg cgg cac acg ccc gag ctg ccc			4446
Ile His Glu Ile Pro Arg Glu Glu Leu Arg His Thr Pro Glu Leu Pro			
1415	1420	1425	1430
ctg gcc ccg cgg ccg ctc aag gag ggc tcc atc acg cag ggc acc ccg			4494
Leu Ala Pro Arg Pro Leu Lys Glu Gly Ser Ile Thr Gln Gly Thr Pro			
1435	1440	1445	
ctc aag tac gac acc ggc gcg tcc acc act ggc tcc aaa aag cac gac			4542
Leu Lys Tyr Asp Thr Gly Ala Ser Thr Thr Gly Ser Lys Lys His Asp			
1450	1455	1460	
gta cgc tcc ctc atc ggc agc ccc ggc cgg acg ttc cca ccc gtg cac			4590
Val Arg Ser Leu Ile Gly Ser Pro Gly Arg Thr Phe Pro Pro Val His			
1465	1470	1475	
ccg ctg gat gtg atg gcc gac gcc cgg gca ctg gaa cgt gcc tgc tac			4638
Pro Leu Asp Val Met Ala Asp Ala Arg Ala Leu Glu Arg Ala Cys Tyr			
1480	1485	1490	
gag gag agc ctg aag agc cgg cca ggg acc gcc agc agc tcg ggg ggc			4686
Glu Glu Ser Leu Lys Ser Arg Pro Gly Thr Ala Ser Ser Ser Gly Gly			
1495	1500	1505	1510
tcc att gcg cgc ggc gcc ccg gtc att gtg cct gag ctg ggt aag ccg			4734
Ser Ile Ala Arg Gly Ala Pro Val Ile Val Pro Glu Leu Gly Lys Pro			
1515	1520	1525	
cgg cag agc ccc ctg acc tat gag gac cac ggg gca ccc ttt gcc ggc			4782
Arg Gln Ser Pro Leu Thr Tyr Glu Asp His Gly Ala Pro Phe Ala Gly			
1530	1535	1540	
cac ctc cca cga ggt tcg ccc gtg acc atg cgg gag ccc acg ccg cgc			4830
His Leu Pro Arg Gly Ser Pro Val Thr Met Arg Glu Pro Thr Pro Arg			
1545	1550	1555	
ctg cag gag ggc agc ctt tcg tcc agc aag gca tcc cag gac cga aag			4878
Leu Gln Glu Gly Ser Leu Ser Ser Ser Lys Ala Ser Gln Asp Arg Lys			

1560	1565	1570	
ctg acg tgc acg cct cgt gag atc gcc aag tcc ccg cac agc acc gtg			4926
Leu Thr Ser Thr Pro Arg Glu Ile Ala Lys Ser Pro His Ser Thr Val			
1575	1580	1585	1590
ccc gag cac cac cca cac ccc atc tgc ccc tat gag cac ctg ctt cgg			4974
Pro Glu His His Pro His Pro Ile Ser Pro Tyr Glu His Leu Leu Arg			
1595	1600		1605
ggc gtg agt ggc gtg gac ctg tat cgc agc cac atc ccc ctg gcc ttc			5022
Gly Val Ser Gly Val Asp Leu Tyr Arg Ser His Ile Pro Leu Ala Phe			
1610	1615		1620
gac ccc acc tcc ata ccc cgc ggc atc cct ctg gac gca gcc gct gcc			5070
Asp Pro Thr Ser Ile Pro Arg Gly Ile Pro Leu Asp Ala Ala Ala Ala			
1625	1630		1635
tac tac ctg ccc cga cac ctg gcc ccc aac ccc acc tac ccg cac ctg			5118
Tyr Tyr Leu Pro Arg His Leu Ala Pro Asn Pro Thr Tyr Pro His Leu			
1640	1645		1650
tac cca ccc tac ctc atc cgc ggc tac ccc gac acg gcg gcg ctg gag			5166
Tyr Pro Pro Tyr Leu Ile Arg Gly Tyr Pro Asp Thr Ala Ala Leu Glu			
1655	1660	1665	1670
aac cgg cag acc atc atc aat gac tac atc acc tgc cag cag atg cac			5214
Asn Arg Gln Thr Ile Ile Asn Asp Tyr Ile Thr Ser Gln Gln Met His			
1675	1680		1685
cac aac acg gcc acc gcc atg gcc cag cga gct gat atg ctg agg ggc			5262
His Asn Thr Ala Thr Ala Met Ala Gln Arg Ala Asp Met Leu Arg Gly			
1690	1695		1700
ctc tgc ccc cgc gag tcc tgc ctg gca ctc aac tac gct gcg ggt ccc			5310
Leu Ser Pro Arg Glu Ser Ser Leu Ala Leu Asn Tyr Ala Ala Gly Pro			
1705	1710		1715
cga ggc atc atc gac ctg tcc caa gtg cca cac ctg cct gtg ctc gtg			5358
Arg Gly Ile Ile Asp Leu Ser Gln Val Pro His Leu Pro Val Leu Val			
1720	1725	1730	
ccc ccg aca cca ggc acc cca gcc acc gcc atg gac cgc ctt gcc tac			5406
Pro Pro Thr Pro Gly Thr Pro Ala Thr Ala Met Asp Arg Leu Ala Tyr			
1735	1740	1745	1750
ctc ccc acc gcg ccc cag ccc ttc agc agc cgc cac agc agc tcc cca			5454
Leu Pro Thr Ala Pro Gln Pro Phe Ser Ser Arg His Ser Ser Ser Pro			
1755	1760		1765
ctc tcc cca gga ggt cca aca cac ttg aca aaa cca acc acc acg tcc			5502
Leu Ser Pro Gly Gly Pro Thr His Leu Thr Lys Pro Thr Thr Thr Ser			
1770	1775		1780
tgc tcc gag cgg gag cga gac cgg gat cga gag cgg gac cgg gat cgg			5550
Ser Ser Glu Arg Glu Arg Asp Arg Asp Arg Glu Arg Asp Arg Asp Arg			
1785	1790	1795	
gag cgg gaa aag tcc atc ctc acg tcc acc acg acg gtg gag cac gca			5598
Glu Arg Glu Lys Ser Ile Leu Thr Ser Thr Thr Thr Val Glu His Ala			
1800	1805	1810	

ccc atc tgg aga cct ggt aca gag cag agc agc ggc agc agc ggc agc Pro Ile Trp Arg Pro Gly Thr Glu Gln Ser Ser Gly Ser Ser Gly Ser 1815 1820 1825 1830	5646
agc ggc ggg ggt ggg ggc agc agc agc cgc ccc gcc tcc cac tcc cat Ser Gly Gly Gly Ser Ser Ser Arg Pro Ala Ser His Ser His 1835 1840 1845	5694
gcc cac cag cac tcg ccc atc tcc cct cgg acc cag gat gcc ctc cag Ala His Gln His Ser Pro Ile Ser Pro Arg Thr Gln Asp Ala Leu Gln 1850 1855 1860	5742
cag aga ccc agt gtg ctt cac aac aca ggc atg aag ggt atc atc acc Gln Arg Pro Ser Val Leu His Asn Thr Gly Met Lys Gly Ile Ile Thr 1865 1870 1875	5790
gct gtg gag ccc agc aag ccc acg gtc ctg agg tcc acc tcc acc tcc Ala Val Glu Pro Ser Lys Pro Thr Val Leu Arg Ser Thr Ser Thr Ser 1880 1885 1890	5838
tca ccc gtt cgc cca gct gcc aca ttc cca cct gcc acc cac tgc cca Ser Pro Val Arg Pro Ala Ala Thr Phe Pro Pro Ala Thr His Cys Pro 1895 1900 1905 1910	5886
ctg ggc ggc acc ctc gat ggg gtc tac cct acc ctc atg gag ccc gtc Leu Gly Gly Thr Leu Asp Gly Val Tyr Pro Thr Leu Met Glu Pro Val 1915 1920 1925	5934
ttg ctg ccc aag gag gcc ccc cgg gtc gcc cgg cca gag cgg ccc cga Leu Leu Pro Lys Glu Ala Pro Arg Val Ala Arg Pro Glu Arg Pro Arg 1930 1935 1940	5982
gca gac acc ggc cat gcc ttc ctc gcc aag ccc cca gcc cgc tcc ggg Ala Asp Thr Gly His Ala Phe Leu Ala Lys Pro Pro Ala Arg Ser Gly 1945 1950 1955	6030
ctg gag ccc gcc tcc tcc ccc agc aag ggc tcg gag ccc cgg ccc cta Leu Glu Pro Ala Ser Ser Pro Ser Lys Gly Ser Glu Pro Arg Pro Leu 1960 1965 1970	6078
gtg cct cct gtc tct ggc cac gcc acc atc gcc cgc acc cct gcg aag Val Pro Pro Val Ser Gly His Ala Thr Ile Ala Arg Thr Pro Ala Lys 1975 1980 1985 1990	6126
aac ctc gca cct cac cac gcc agc ccg gac ccg ccg gcg cca cct gcc Asn Leu Ala Pro His His Ala Ser Pro Asp Pro Pro Ala Pro Pro Ala 1995 2000 2005	6174
tcg gcc tcg gac ccg cac cgg gaa aag act caa agt aaa ccc ttt tcc Ser Ala Ser Asp Pro His Arg Glu Lys Thr Gln Ser Lys Pro Phe Ser 2010 2015 2020	6222
atc cag gaa ctg gaa ctc cgt tct ctg ggt tac cac ggc agc agc tac Ile Gln Glu Leu Glu Leu Arg Ser Leu Gly Tyr His Gly Ser Ser Tyr 2025 2030 2035	6270
agc ccc gaa ggg gtg gag ccc gtc agc cct gtg agc tca ccc agt ctg Ser Pro Glu Gly Val Glu Pro Val Ser Pro Val Ser Ser Pro Ser Leu 2040 2045 2050	6318

Asn Ile Ser Gln Pro Gly Thr Glu Ile Phe Asn Met Pro Ala Ile Thr
 2295 2300 2305 2310

gga aca ggc ctt atg acc tat aga agc cag gcg gtg cag gaa cat gcc 7134
 Gly Thr Gly Leu Met Thr Tyr Arg Ser Gln Ala Val Gln Glu His Ala
 2315 2320 2325

agc acc aac atg ggg ctg gag gcc ata att aga aag gca ctc atg ggt 7182
 Ser Thr Asn Met Gly Leu Glu Ala Ile Ile Arg Lys Ala Leu Met Gly
 2330 2335 2340

aaa tat gac cag tgg gaa gag tcc ccg ccg ctc agc gcc aat gct ttt 7230
 Lys Tyr Asp Gln Trp Glu Glu Ser Pro Pro Leu Ser Ala Asn Ala Phe
 2345 2350 2355

aac cct ctg aat gcc agt gcc agc ctg ccc gct gct atg ccc ata acc 7278
 Asn Pro Leu Asn Ala Ser Ala Ser Leu Pro Ala Ala Met Pro Ile Thr
 2360 2365 2370

gct gct gac gga cgg agt gac cac aca ctc acc tcg cca ggt ggc ggc 7326
 Ala Ala Asp Gly Arg Ser Asp His Thr Leu Thr Ser Pro Gly Gly Gly
 2375 2380 2385 2390

ggg aag gcc aag gtc tct ggc aga ccc agc agc cga aaa gcc aag tcc 7374
 Gly Lys Ala Lys Val Ser Gly Arg Pro Ser Ser Arg Lys Ala Lys Ser
 2395 2400 2405

ccg gcc ccg ggc ctg gca tct ggg gac cgg cca ccc tct gtc tcc tca 7422
 Pro Ala Pro Gly Leu Ala Ser Gly Asp Arg Pro Pro Ser Val Ser Ser
 2410 2415 2420

gtg cac tcg gag gga gac tgc aac cgc cgg acg ccg ctc acc aac cgc 7470
 Val His Ser Glu Gly Asp Cys Asn Arg Arg Thr Pro Leu Thr Asn Arg
 2425 2430 2435

gtg tgg gag gac agg ccc tcg tcc gca ggt tcc acg cca ttc ccc tac 7518
 Val Trp Glu Asp Arg Pro Ser Ser Ala Gly Ser Thr Pro Phe Pro Tyr
 2440 2445 2450

aac ccc ctg atc atg cgg ctg cag gcg ggt gtc atg gct tcc cca ccc 7566
 Asn Pro Leu Ile Met Arg Leu Gln Ala Gly Val Met Ala Ser Pro Pro
 2455 2460 2465 2470

cca ccg ggc ctc ccc gcg ggc agc ggg ccc ctc gct ggc ccc cac cac 7614
 Pro Pro Gly Leu Pro Ala Gly Ser Gly Pro Leu Ala Gly Pro His His
 2475 2480 2485

gcc tgg gac gag gag ccc aag cca ctg ctc tgc tcg cag tac gag aca 7662
 Ala Trp Asp Glu Glu Pro Lys Pro Leu Leu Cys Ser Gln Tyr Glu Thr
 2490 2495 2500

ctc tcc gac agc gag tgactcagaa cagggcgggg gggggcgggc ggtgtcaggt 7717
 Leu Ser Asp Ser Glu
 2505

cccagcgagc cacaggaacg gccctgcagg agcgggggcg ctgccgactc ccccaaccaa 7777

ggaaggagcc cctgagtccg cctgcgcctc catccatctg tccgtccaga gccggcatcc 7837

ttgcctgtct aaagccttaa ctaagactcc cgccccgggc tggccctgtg cagaccttac 7897

CCCTGCTCT

```

<210> 2
<211> 2507
<212> PRT
<213> Homo sapiens

<400> 2
Met Ser Gly Ser Thr Gln Pro Val Ala Gln Thr Trp Arg Ala Thr Glu
  1                               5                10              15
Pro Arg Tyr Pro Pro His Ser Leu Ser Tyr Pro Val Gln Ile Ala Arg
                20                        25              30
Thr His Thr Asp Val Gly Leu Leu Glu Tyr Gln His His Ser Arg Asp
          35                      40              45
Tyr Ala Ser His Leu Ser Pro Gly Ser Ile Ile Gln Pro Gln Arg Arg
      50                      55              60
Arg Pro Ser Leu Leu Ser Glu Phe Gln Pro Gly Asn Glu Arg Ser Gln
  65                      70              75              80
Glu Leu His Leu Arg Pro Glu Ser His Ser Tyr Leu Pro Glu Leu Gly
                      85              90              95
Lys Ser Glu Met Glu Phe Ile Glu Ser Lys Arg Pro Arg Leu Glu Leu
          100                      105              110
Leu Pro Asp Pro Leu Leu Arg Pro Ser Pro Leu Leu Ala Thr Gly Gln
          115                      120              125
Pro Ala Gly Ser Glu Asp Leu Thr Lys Asp Arg Ser Leu Thr Gly Lys
      130                      135              140

```

Leu	Glu	Pro	Val	Ser	Pro	Pro	Ser	Pro	Pro	His	Thr	Asp	Pro	Glu	Leu	145	150	155	160
Glu	Leu	Val	Pro	Pro	Arg	Leu	Ser	Lys	Glu	Glu	Leu	Ile	Gln	Asn	Met	165	170	175	
Asp	Arg	Val	Asp	Arg	Glu	Ile	Thr	Met	Val	Glu	Gln	Gln	Ile	Ser	Lys	180	185	190	
Leu	Lys	Lys	Lys	Gln	Gln	Gln	Leu	Glu	Glu	Glu	Ala	Ala	Lys	Pro	Pro	195	200	205	
Glu	Pro	Glu	Lys	Pro	Val	Ser	Pro	Pro	Pro	Ile	Glu	Ser	Lys	His	Arg	210	215	220	
Ser	Leu	Val	Gln	Ile	Ile	Tyr	Asp	Glu	Asn	Arg	Lys	Lys	Ala	Glu	Ala	225	230	235	240
Ala	His	Arg	Ile	Leu	Glu	Gly	Leu	Gly	Pro	Gln	Val	Glu	Leu	Pro	Leu	245	250	255	
Tyr	Asn	Gln	Pro	Ser	Asp	Thr	Arg	Gln	Tyr	His	Glu	Asn	Ile	Lys	Ile	260	265	270	
Asn	Gln	Ala	Met	Arg	Lys	Lys	Leu	Ile	Leu	Tyr	Phe	Lys	Arg	Arg	Asn	275	280	285	
His	Ala	Arg	Lys	Gln	Trp	Glu	Gln	Lys	Phe	Cys	Gln	Arg	Tyr	Asp	Gln	290	295	300	
Leu	Met	Glu	Ala	Trp	Glu	Lys	Lys	Val	Glu	Arg	Ile	Glu	Asn	Asn	Pro	305	310	315	320
Arg	Arg	Arg	Ala	Lys	Glu	Ser	Lys	Val	Arg	Glu	Tyr	Tyr	Glu	Lys	Gln	325	330	335	
Phe	Pro	Glu	Ile	Arg	Lys	Gln	Arg	Glu	Leu	Gln	Glu	Arg	Met	Gln	Arg	340	345	350	
Val	Gly	Gln	Arg	Gly	Ser	Gly	Leu	Ser	Met	Ser	Pro	Ala	Arg	Ser	Glu	355	360	365	
His	Glu	Val	Ser	Glu	Ile	Ile	Asp	Gly	Leu	Ser	Glu	Gln	Glu	Asn	Leu	370	375	380	
Glu	Lys	Gln	Met	Arg	Gln	Leu	Ala	Val	Ile	Pro	Pro	Met	Leu	Tyr	Asp	385	390	395	400
Ala	Asp	Gln	Gln	Arg	Ile	Lys	Phe	Ile	Asn	Met	Asn	Gly	Leu	Met	Ala	405	410	415	
Asp	Pro	Met	Lys	Val	Tyr	Lys	Asp	Arg	Gln	Val	Met	Asn	Met	Trp	Ser	420	425	430	
Glu	Gln	Glu	Lys	Glu	Thr	Phe	Arg	Glu	Lys	Phe	Met	Gln	His	Pro	Lys	435	440	445	
Asn	Phe	Gly	Leu	Ile	Ala	Ser	Phe	Leu	Glu	Arg	Lys	Thr	Val	Ala	Glu	450	455	460	

10250-10250

Cys 465	Val	Leu	Tyr	Tyr	Tyr 470	Leu	Thr	Lys	Lys	Asn 475	Glu	Asn	Tyr	Lys	Ser 480
Leu	Val	Arg	Arg	Ser 485	Tyr	Arg	Arg	Arg	Gly 490	Lys	Ser	Gln	Gln	Gln 495	Gln
Gln	Gln	Gln	Gln	Gln 500	Gln	Gln	Gln	Gln	Gln 505	Gln	Gln	Gln	Pro 510	Met	Pro
Arg	Ser	Ser 515	Gln	Glu	Glu	Lys	Asp 520	Glu	Lys	Glu	Lys	Glu 525	Lys	Glu	Ala
Glu	Lys 530	Glu	Glu	Glu	Lys	Pro 535	Glu	Val	Glu	Asn	Asp 540	Lys	Glu	Asp	Leu
Leu 545	Lys	Glu	Lys	Thr	Asp 550	Asp	Thr	Ser	Gly	Glu 555	Asp	Asn	Asp	Glu	Lys 560
Glu	Ala	Val	Ala	Ser 565	Lys	Gly	Arg	Lys	Thr 570	Ala	Asn	Ser	Gln	Gly 575	Arg
Arg	Lys	Gly	Arg 580	Ile	Thr	Arg	Ser	Met 585	Ala	Asn	Glu	Ala	Asn 590	Ser	Glu
Glu	Ala	Ile 595	Thr	Pro	Gln	Gln	Ser 600	Ala	Glu	Leu	Ala	Ser 605	Met	Glu	Leu
Asn 610	Glu	Ser	Ser	Arg	Trp	Thr 615	Glu	Glu	Glu	Met	Glu 620	Thr	Ala	Lys	Lys
Gly 625	Leu	Leu	Glu	His	Gly 630	Arg	Asn	Trp	Ser	Ala 635	Ile	Ala	Arg	Met	Val 640
Gly	Ser	Lys	Thr 645	Val	Ser	Gln	Cys	Lys	Asn 650	Phe	Tyr	Phe	Asn	Tyr 655	Lys
Lys	Arg	Gln	Asn 660	Leu	Asp	Glu	Ile	Leu 665	Gln	Gln	His	Lys	Leu 670	Lys	Met
Glu	Lys	Glu 675	Arg	Asn	Ala	Arg	Arg 680	Lys	Lys	Lys	Lys	Ala 685	Pro	Ala	Ala
Ala 690	Ser	Glu	Glu	Ala	Ala	Phe 695	Pro	Pro	Val	Val	Glu 700	Asp	Glu	Glu	Met
Glu 705	Ala	Ser	Gly	Val	Thr 710	Gly	Asn	Glu	Glu	Glu 715	Met	Val	Glu	Glu	Ala 720
Glu	Ala	Thr	Val 725	Asn	Asn	Ser	Ser	Asp 730	Thr	Glu	Ser	Ile	Pro	Ser 735	Pro
His	Thr	Glu 740	Ala	Ala	Lys	Asp	Thr	Gly 745	Gln	Asn	Gly	Pro	Lys 750	Pro	Pro
Ala	Thr	Leu 755	Gly	Ala	Asp	Gly	Pro 760	Pro	Pro	Gly	Pro	Pro 765	Thr	Pro	Pro
Pro 770	Glu	Asp	Ile	Pro	Ala	Pro 775	Thr	Glu	Ser	Thr	Pro 780	Ala	Ser	Glu	Ala
Thr	Leu	Ala	Pro	Thr	Pro	Pro	Pro	Ala	Pro	Pro	Phe	Pro	Ser	Ser	Pro

785		790		795		800
Pro Pro Val Val	Pro Lys Glu Glu Lys	Glu Glu Glu Thr	Ala Ala Ala			
	805		810		815	
Pro Pro Val Glu	Glu Gly Glu Glu Gln	Lys Pro Pro Ala	Ala Glu Glu			
	820		825		830	
Leu Ala Val Asp	Thr Gly Lys Ala Glu	Glu Glu Pro Val	Lys Ser Glu Cys			
	835		840		845	
Thr Glu Glu Ala	Glu Glu Gly Pro Ala	Lys Gly Lys Asp	Ala Glu Ala			
	850		855		860	
Ala Glu Ala Thr	Ala Glu Arg Ala Leu	Lys Ala Glu Lys	Lys Glu Gly			
	865		870		875	
Gly Ser Gly Arg	Ala Thr Thr Ala Lys	Ser Ser Gly Ala	Pro Gln Asp			
	885		890		895	
Ser Asp Ser Ser	Ala Thr Cys Ser Ala	Asp Glu Val Asp	Glu Ala Glu			
	900		905		910	
Gly Gly Asp Lys	Asn Arg Leu Leu Ser	Pro Arg Pro Ser	Leu Leu Thr			
	915		920		925	
Pro Thr Gly Asp	Pro Arg Ala Asn Ala	Ser Pro Gln Lys	Pro Leu Asp			
	930		935		940	
Leu Lys Gln Leu	Lys Gln Arg Ala Ala	Ala Ile Pro Pro	Ile Gln Val			
	945		950		955	
Thr Lys Val His	Glu Pro Pro Arg Glu	Asp Ala Ala Pro	Thr Lys Pro			
	965		970		975	
Ala Pro Pro Ala	Pro Pro Pro Pro Gln	Asn Leu Gln Pro	Glu Ser Asp			
	980		985		990	
Ala Pro Gln Gln	Pro Gly Ser Ser Pro	Arg Gly Lys Ser	Arg Ser Pro			
	995		1000		1005	
Ala Pro Pro Ala	Asp Lys Glu Ala Glu	Lys Pro Val Phe	Phe Pro Ala			
	1010		1015		1020	
Phe Ala Ala Glu	Ala Gln Lys Leu Pro	Gly Asp Pro Pro	Cys Trp Thr			
	1025		1030		1035	
Ser Gly Leu Pro	Phe Pro Val Pro Pro	Arg Glu Val Ile	Lys Ala Ser			
	1045		1050		1055	
Pro His Ala Pro	Asp Pro Ser Ala Phe	Ser Tyr Ala Pro	Pro Gly His			
	1060		1065		1070	
Pro Leu Pro Leu	Gly Leu His Asp Thr	Ala Arg Pro Val	Leu Pro Arg			
	1075		1080		1085	
Pro Pro Thr Ile	Ser Asn Pro Pro Pro	Leu Ile Ser Ser	Ala Lys His			
	1090		1095		1100	
Pro Ser Val Leu	Glu Arg Gln Ile Gly	Ala Ile Ser Gln	Gly Met Ser			
	1105		1110		1115	
					1120	

102220-102220

Val	Gln	Leu	His	Val	Pro	Tyr	Ser	Glu	His	Ala	Lys	Ala	Pro	Val	Gly	
				1125						1130					1135	
Pro	Val	Thr	Met	Gly	Leu	Pro	Leu	Pro	Met	Asp	Pro	Lys	Lys	Leu	Ala	
			1140					1145					1150			
Pro	Phe	Ser	Gly	Val	Lys	Gln	Glu	Gln	Leu	Ser	Pro	Arg	Gly	Gln	Ala	
		1155					1160					1165				
Gly	Pro	Pro	Glu	Ser	Leu	Gly	Val	Pro	Thr	Ala	Gln	Glu	Ala	Ser	Val	
			1170				1175					1180				
Leu	Arg	Gly	Thr	Ala	Leu	Gly	Ser	Val	Pro	Gly	Gly	Ser	Ile	Thr	Lys	
185					1190					1195					1200	
Gly	Ile	Pro	Ser	Thr	Arg	Val	Pro	Ser	Asp	Ser	Ala	Ile	Thr	Tyr	Arg	
				1205					1210						1215	
Gly	Ser	Ile	Thr	His	Gly	Thr	Pro	Ala	Asp	Val	Leu	Tyr	Lys	Gly	Thr	
			1220					1225						1230		
Ile	Thr	Arg	Ile	Ile	Gly	Glu	Asp	Ser	Pro	Ser	Arg	Leu	Asp	Arg	Gly	
		1235					1240					1245				
Arg	Glu	Asp	Ser	Leu	Pro	Lys	Gly	His	Val	Ile	Tyr	Glu	Gly	Lys	Lys	
	1250						1255				1260					
Gly	His	Val	Leu	Ser	Tyr	Glu	Gly	Gly	Met	Ser	Val	Thr	Gln	Cys	Ser	
265					1270					1275					1280	
Lys	Glu	Asp	Gly	Arg	Ser	Ser	Ser	Gly	Pro	Pro	His	Glu	Thr	Ala	Ala	
				1285					1290					1295		
Pro	Lys	Arg	Thr	Tyr	Asp	Met	Met	Glu	Gly	Arg	Val	Gly	Arg	Ala	Ile	
			1300					1305					1310			
Ser	Ser	Ala	Ser	Ile	Glu	Gly	Leu	Met	Gly	Arg	Ala	Ile	Pro	Pro	Glu	
		1315					1320					1325				
Arg	His	Ser	Pro	His	His	Leu	Lys	Glu	Gln	His	His	Ile	Arg	Gly	Ser	
						1335					1340					
Ile	Thr	Gln	Gly	Ile	Pro	Arg	Ser	Tyr	Val	Glu	Ala	Gln	Glu	Asp	Tyr	
345					1350					1355					1360	
Leu	Arg	Arg	Glu	Ala	Lys	Leu	Leu	Lys	Arg	Glu	Gly	Thr	Pro	Pro	Pro	
				1365					1370					1375		
Pro	Pro	Pro	Ser	Arg	Asp	Leu	Thr	Glu	Ala	Tyr	Lys	Thr	Gln	Ala	Leu	
				1380				1385					1390			
Gly	Pro	Leu	Lys	Leu	Lys	Pro	Ala	His	Glu	Gly	Leu	Val	Ala	Thr	Val	
			1395				1400					1405				
Lys	Glu	Ala	Gly	Arg	Ser	Ile	His	Glu	Ile	Pro	Arg	Glu	Glu	Leu	Arg	
						1415					1420					
His	Thr	Pro	Glu	Leu	Pro	Leu	Ala	Pro	Arg	Pro	Leu	Lys	Glu	Gly	Ser	
425					1430						1435				1440	

Ile Thr Gln Gly Thr Pro Leu Lys Tyr Asp Thr Gly Ala Ser Thr Thr
1445 1450 1455

Gly Ser Lys Lys His Asp Val Arg Ser Leu Ile Gly Ser Pro Gly Arg
1460 1465 1470

Thr Phe Pro Pro Val His Pro Leu Asp Val Met Ala Asp Ala Arg Ala
1475 1480 1485

Leu Glu Arg Ala Cys Tyr Glu Glu Ser Leu Lys Ser Arg Pro Gly Thr
1490 1495 1500

Ala Ser Ser Ser Gly Gly Ser Ile Ala Arg Gly Ala Pro Val Ile Val
505 1510 1515 1520

Pro Glu Leu Gly Lys Pro Arg Gln Ser Pro Leu Thr Tyr Glu Asp His
1525 1530 1535

Gly Ala Pro Phe Ala Gly His Leu Pro Arg Gly Ser Pro Val Thr Met
1540 1545 1550

Arg Glu Pro Thr Pro Arg Leu Gln Glu Gly Ser Leu Ser Ser Ser Lys
1555 1560 1565

Ala Ser Gln Asp Arg Lys Leu Thr Ser Thr Pro Arg Glu Ile Ala Lys
1570 1575 1580

Ser Pro His Ser Thr Val Pro Glu His His Pro His Pro Ile Ser Pro
585 1590 1595 1600

Tyr Glu His Leu Leu Arg Gly Val Ser Gly Val Asp Leu Tyr Arg Ser
1605 1610 1615

His Ile Pro Leu Ala Phe Asp Pro Thr Ser Ile Pro Arg Gly Ile Pro
1620 1625 1630

Leu Asp Ala Ala Ala Ala Tyr Tyr Leu Pro Arg His Leu Ala Pro Asn
1635 1640 1645

Pro Thr Tyr Pro His Leu Tyr Pro Pro Tyr Leu Ile Arg Gly Tyr Pro
1650 1655 1660

Asp Thr Ala Ala Leu Glu Asn Arg Gln Thr Ile Ile Asn Asp Tyr Ile
665 1670 1675 1680

Thr Ser Gln Gln Met His His Asn Thr Ala Thr Ala Met Ala Gln Arg
1685 1690 1695

Ala Asp Met Leu Arg Gly Leu Ser Pro Arg Glu Ser Ser Leu Ala Leu
1700 1705 1710

Asn Tyr Ala Ala Gly Pro Arg Gly Ile Ile Asp Leu Ser Gln Val Pro
1715 1720 1725

His Leu Pro Val Leu Val Pro Pro Thr Pro Gly Thr Pro Ala Thr Ala
1730 1735 1740

Met Asp Arg Leu Ala Tyr Leu Pro Thr Ala Pro Gln Pro Phe Ser Ser
745 1750 1755 1760

Arg His Ser Ser Ser Pro Leu Ser Pro Gly Gly Pro Thr His Leu Thr

1445-1455-1460-1470-1475-1480-1485-1490-1495-1500-1505-1510-1515-1520-1525-1530-1535-1540-1545-1550-1555-1560-1565-1570-1575-1580-1585-1590-1595-1600-1605-1610-1615-1620-1625-1630-1635-1640-1645-1650-1655-1660-1665-1670-1675-1680-1685-1690-1695-1700-1705-1710-1715-1720-1725-1730-1735-1740-1745-1750-1755-1760

Lys	Pro	Thr	Thr	Thr	Ser	Ser	Ser	Glu	Arg	Glu	Arg	Asp	Arg	Asp	Arg	
			1780					1785					1790			
Glu	Arg	Asp	Arg	Asp	Arg	Glu	Arg	Glu	Lys	Ser	Ile	Leu	Thr	Ser	Thr	
		1795					1800					1805				
Thr	Thr	Val	Glu	His	Ala	Pro	Ile	Trp	Arg	Pro	Gly	Thr	Glu	Gln	Ser	
		1810					1815				1820					
Ser	Gly	Ser	Ser	Gly	Ser	Ser	Gly	Gly	Gly	Gly	Gly	Ser	Ser	Ser	Arg	
825					1830					1835					1840	
Pro	Ala	Ser	His	Ser	His	Ala	His	Gln	His	Ser	Pro	Ile	Ser	Pro	Arg	
				1845					1850					1855		
Thr	Gln	Asp	Ala	Leu	Gln	Gln	Arg	Pro	Ser	Val	Leu	His	Asn	Thr	Gly	
			1860					1865					1870			
Met	Lys	Gly	Ile	Ile	Thr	Ala	Val	Glu	Pro	Ser	Lys	Pro	Thr	Val	Leu	
		1875					1880					1885				
Arg	Ser	Thr	Ser	Thr	Ser	Ser	Pro	Val	Arg	Pro	Ala	Ala	Thr	Phe	Pro	
		1890					1895				1900					
Pro	Ala	Thr	His	Cys	Pro	Leu	Gly	Gly	Thr	Leu	Asp	Gly	Val	Tyr	Pro	
905					1910					1915					1920	
Thr	Leu	Met	Glu	Pro	Val	Leu	Leu	Pro	Lys	Glu	Ala	Pro	Arg	Val	Ala	
				1925					1930					1935		
Arg	Pro	Glu	Arg	Pro	Arg	Ala	Asp	Thr	Gly	His	Ala	Phe	Leu	Ala	Lys	
			1940					1945					1950			
Pro	Pro	Ala	Arg	Ser	Gly	Leu	Glu	Pro	Ala	Ser	Ser	Pro	Ser	Lys	Gly	
		1955					1960					1965				
Ser	Glu	Pro	Arg	Pro	Leu	Val	Pro	Pro	Val	Ser	Gly	His	Ala	Thr	Ile	
		1970					1975				1980					
Ala	Arg	Thr	Pro	Ala	Lys	Asn	Leu	Ala	Pro	His	His	Ala	Ser	Pro	Asp	
985					1990					1995					2000	
Pro	Pro	Ala	Pro	Pro	Ala	Ser	Ala	Ser	Asp	Pro	His	Arg	Glu	Lys	Thr	
				2005					2010					2015		
Gln	Ser	Lys	Pro	Phe	Ser	Ile	Gln	Glu	Leu	Glu	Leu	Arg	Ser	Leu	Gly	
			2020					2025					2030			
Tyr	His	Gly	Ser	Ser	Tyr	Ser	Pro	Glu	Gly	Val	Glu	Pro	Val	Ser	Pro	
		2035					2040					2045				
Val	Ser	Ser	Pro	Ser	Leu	Thr	His	Asp	Lys	Gly	Leu	Pro	Lys	His	Leu	
		2050					2055				2060					
Glu	Glu	Leu	Asp	Lys	Ser	His	Leu	Glu	Gly	Glu	Leu	Arg	Pro	Lys	Gln	
065					2070					2075					2080	
Pro	Gly	Pro	Val	Lys	Leu	Gly	Gly	Glu	Ala	Ala	His	Leu	Pro	His	Leu	
				2085					2090					2095		

Arg	Pro	Leu	Pro	Glu	Ser	Gln	Pro	Ser	Ser	Ser	Pro	Leu	Gln	Thr	
				2100						2105					2110
Ala	Pro	Gly	Val	Lys	Gly	His	Gln	Arg	Val	Val	Thr	Leu	Ala	Gln	His
		2115					2120					2125			
Ile	Ser	Glu	Val	Ile	Thr	Gln	Asp	Tyr	Thr	Arg	His	His	Pro	Gln	Gln
	2130					2135					2140				
Leu	Ser	Ala	Pro	Leu	Pro	Ala	Pro	Leu	Tyr	Ser	Phe	Pro	Gly	Ala	Ser
145					2150					2155					2160
Cys	Pro	Val	Leu	Asp	Leu	Arg	Arg	Pro	Pro	Ser	Asp	Leu	Tyr	Leu	Pro
			2165						2170					2175	
Pro	Pro	Asp	His	Gly	Ala	Pro	Ala	Arg	Gly	Ser	Pro	His	Ser	Glu	Gly
		2180					2185						2190		
Gly	Lys	Arg	Ser	Pro	Glu	Pro	Asn	Lys	Thr	Ser	Val	Leu	Gly	Gly	Gly
	2195						2200					2205			
Glu	Asp	Gly	Ile	Glu	Pro	Val	Ser	Pro	Pro	Glu	Gly	Met	Thr	Glu	Pro
	2210					2215					2220				
Gly	His	Ser	Arg	Ser	Ala	Val	Tyr	Pro	Leu	Leu	Tyr	Arg	Asp	Gly	Glu
225					2230					2235					2240
Gln	Thr	Glu	Pro	Ser	Arg	Met	Gly	Ser	Lys	Ser	Pro	Gly	Asn	Thr	Ser
			2245						2250					2255	
Gln	Pro	Pro	Ala	Phe	Phe	Ser	Lys	Leu	Thr	Glu	Ser	Asn	Ser	Ala	Met
		2260					2265						2270		
Val	Lys	Ser	Lys	Lys	Gln	Glu	Ile	Asn	Lys	Lys	Leu	Asn	Thr	His	Asn
	2275					2280						2285			
Arg	Asn	Glu	Pro	Glu	Tyr	Asn	Ile	Ser	Gln	Pro	Gly	Thr	Glu	Ile	Phe
	2290					2295					2300				
Asn	Met	Pro	Ala	Ile	Thr	Gly	Thr	Gly	Leu	Met	Thr	Tyr	Arg	Ser	Gln
305				2310						2315					2320
Ala	Val	Gln	Glu	His	Ala	Ser	Thr	Asn	Met	Gly	Leu	Glu	Ala	Ile	Ile
			2325						2330					2335	
Arg	Lys	Ala	Leu	Met	Gly	Lys	Tyr	Asp	Gln	Trp	Glu	Glu	Ser	Pro	Pro
		2340						2345					2350		
Leu	Ser	Ala	Asn	Ala	Phe	Asn	Pro	Leu	Asn	Ala	Ser	Ala	Ser	Leu	Pro
		2355				2360						2365			
Ala	Ala	Met	Pro	Ile	Thr	Ala	Ala	Asp	Gly	Arg	Ser	Asp	His	Thr	Leu
	2370					2375					2380				
Thr	Ser	Pro	Gly	Gly	Gly	Gly	Lys	Ala	Lys	Val	Ser	Gly	Arg	Pro	Ser
385				2390						2395					2400
Ser	Arg	Lys	Ala	Lys	Ser	Pro	Ala	Pro	Gly	Leu	Ala	Ser	Gly	Asp	Arg
			2405						2410					2415	

Pro Pro Ser Val Ser Ser Val His Ser Glu Gly Asp Cys Asn Arg Arg
 2420 2425 2430

Thr Pro Leu Thr Asn Arg Val Trp Glu Asp Arg Pro Ser Ser Ala Gly
 2435 2440 2445

Ser Thr Pro Phe Pro Tyr Asn Pro Leu Ile Met Arg Leu Gln Ala Gly
 2450 2455 2460

Val Met Ala Ser Pro Pro Pro Pro Gly Leu Pro Ala Gly Ser Gly Pro
 465 2470 2475 2480

Leu Ala Gly Pro His His Ala Trp Asp Glu Glu Pro Lys Pro Leu Leu
 2485 2490 2495

Cys Ser Gln Tyr Glu Thr Leu Ser Asp Ser Glu
 2500 2505

<210> 3
 <211> 7521
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(7521)

<400> 3
 atg tcg ggc tcc aca cag cct gtg gca cag acg tgg agg gcc act gag 48
 Met Ser Gly Ser Thr Gln Pro Val Ala Gln Thr Trp Arg Ala Thr Glu
 1 5 10 15

ccc cgc tac ccg ccc cac agc ctt tcc tac cca gtg cag atc gcc cgg 96
 Pro Arg Tyr Pro Pro His Ser Leu Ser Tyr Pro Val Gln Ile Ala Arg
 20 25 30

acg cac acg gac gtc ggg ctc ctg gag tac cag cac cac tcc cgc gac 144
 Thr His Thr Asp Val Gly Leu Leu Glu Tyr Gln His His Ser Arg Asp
 35 40 45

tat gcc tcc cac ctg tcg ccc ggc tcc atc atc cag ccc cag cgg cgg 192
 Tyr Ala Ser His Leu Ser Pro Gly Ser Ile Ile Gln Pro Gln Arg Arg
 50 55 60

agg ccc tcc ctg ctg tct gag ttc cag ccc ggg aat gaa cgg tcc cag 240
 Arg Pro Ser Leu Leu Ser Glu Phe Gln Pro Gly Asn Glu Arg Ser Gln
 65 70 75 80

gag ctc cac ctg cgg cca gag tcc cac tca tac ctg ccc gag ctg ggg 288
 Glu Leu His Leu Arg Pro Glu Ser His Ser Tyr Leu Pro Glu Leu Gly
 85 90 95

aag tca gag atg gag ttc att gaa agc aag cgc cct cgg cta gag ctg 336
 Lys Ser Glu Met Glu Phe Ile Glu Ser Lys Arg Pro Arg Leu Glu Leu
 100 105 110

ctg cct gac ccc ctg ctg cga ccg tca ccc ctg ctg gcc acg ggc cag 384
 Leu Pro Asp Pro Leu Leu Arg Pro Ser Pro Leu Leu Ala Thr Gly Gln
 115 120 125

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cct Pro	gcg Ala	gga Gly	tct Ser	gaa Glu	gac Asp	ctc Leu	acc Thr	aag Lys	gac Asp	cgt Arg	agc Ser	ctg Leu	acg Thr	ggc Gly	aag Lys	432
130135140																
ctg Leu	gaa Glu	ccg Pro	gtg Val	tct Ser	ccc Pro	ccc Pro	agc Ser	ccc Pro	ccg Pro	cac His	act Thr	gac Asp	cct Pro	gag Glu	ctg Leu	480
145150155160																
gag Glu	ctg Leu	gtg Val	ccg Pro	cca Pro	cgg Arg	ctg Leu	tcc Ser	aag Lys	gag Glu	gag Glu	ctg Leu	atc Ile	cag Gln	aac Asn	atg Met	528
165170175																
gac Asp	cgc Arg	gtg Val	gac Asp	cga Arg	gag Glu	atc Ile	acc Thr	atg Met	gta Val	gag Glu	cag Gln	cag Gln	atc Ile	tct Ser	aag Lys	576
180185190																
ctg Leu	aag Lys	aag Lys	aag Lys	cag Gln	caa Gln	cag Gln	ctg Leu	gag Glu	gag Glu	gag Glu	gct Ala	gcc Ala	aag Lys	ccg Pro	ccc Pro	624
195200205																
gag Glu	cct Pro	gag Glu	aag Lys	ccc Pro	gtg Val	tca Ser	ccg Pro	ccg Pro	ccc Pro	atc Ile	gag Glu	tcg Ser	aag Lys	cac His	cgc Arg	672
210215220																
agc Ser	ctg Leu	gtg Val	cag Gln	atc Ile	atc Ile	tac Tyr	gac Asp	gag Glu	aac Asn	cgg Arg	aag Lys	aag Lys	gct Ala	gaa Glu	gct Ala	720
225230235240																
gca Ala	cat His	cgg Arg	att Ile	ctg Leu	gaa Glu	ggc Gly	ctg Leu	ggg Gly	ccc Pro	cag Gln	gtg Val	gag Glu	ctg Leu	ccg Pro	ctg Leu	768
245250255																
tac Tyr	aac Asn	cag Gln	ccc Pro	tcc Ser	gac Asp	acc Thr	cgg Arg	cag Gln	tat Tyr	cat His	gag Glu	aac Asn	atc Ile	aaa Lys	ata Ile	816
260265270																
aac Asn	cag Gln	gcg Ala	atg Met	cgg Arg	aag Lys	aag Lys	cta Leu	atc Ile	ttg Leu	tac Tyr	ttc Phe	aag Lys	agg Arg	agg Arg	aat Asn	864
275280285																
cac His	gct Ala	cgg Arg	aaa Lys	caa Gln	tgg Trp	gag Glu	cag Gln	aag Lys	ttc Phe	tgc Cys	cag Gln	cgc Arg	tat Tyr	gac Asp	cag Gln	912
290295300																
ctc Leu	atg Met	gag Glu	gcc Ala	tgg Trp	gag Glu	aag Lys	aag Lys	gtg Val	gag Glu	cgc Arg	atc Ile	gag Glu	aac Asn	aac Asn	ccc Pro	960
305310315320																
cgg Arg	cgg Arg	cgg Arg	gcc Ala	aag Lys	gag Glu	agc Ser	aag Lys	gtt Val	cgc Arg	gag Glu	tac Tyr	tac Tyr	gag Glu	aag Lys	cag Gln	1008
325330335																
ttc Phe	cct Pro	gag Glu	atc Ile	cgc Arg	aag Lys	cag Gln	cgc Arg	gag Glu	ctg Leu	cag Gln	gag Glu	cgc Arg	atg Met	cag Gln	agg Arg	1056
340345350																
gtg Val	ggc Gly	cag Gln	cgg Arg	ggc Gly	agt Ser	ggg Gly	ctg Leu	tcc Ser	atg Met	tcg Ser	ccc Pro	gcc Ala	cgc Arg	agc Ser	gag Glu	1104
355360365																
cac Glu	gag Glu	gtg Val	tca Val	gag Glu	atc Ile	atc Ile	gat Asp	ggc Gly	ctc Leu	tca Val	gag Glu	cag Glu	gag Glu	aac Asn	ctg Glu	1152

His 370	Glu	Val	Ser	Glu	Ile	Ile	Asp	Gly	Leu	Ser	Glu	Gln	Glu	Asn	Leu	
						375					380					
gag Glu 385	aag Lys	cag Gln	atg Met	cgc Arg	cag Gln	ctg Leu	gcc Ala	gtg Val	atc Ile	ccg Pro	ccc Pro	atg Met	ctg Leu	tac Tyr	gac Asp	1200 400
gct Ala	gac Asp	cag Gln	cag Gln	cgc Arg	atc Ile	aag Lys	ttc Phe	atc Ile	aac Asn	atg Met	aac Asn	ggg Gly	ctt Leu	atg Met	gcc Ala	1248 415
gac Asp	ccc Pro	atg Met	aag Lys	gtg Val	tac Tyr	aaa Lys	gac Asp	cgc Arg	cag Gln	gtc Val	atg Met	aac Asn	atg Met	tgg Trp	agt Ser	1296 430
gag Glu	cag Gln	gag Glu	aag Lys	gag Glu	acc Thr	ttc Phe	cgg Arg	gag Glu	aag Lys	ttc Phe	atg Met	cag Gln	cat His	ccc Pro	aag Lys	1344 445
aac Asn	ttt Phe	ggc Gly	ctg Leu	atc Ile	gca Ala	tca Ser	ttc Phe	ctg Leu	gag Glu	agg Arg	aag Lys	aca Thr	gtg Val	gct Ala	gag Glu	1392 460
tgc Cys 465	gtc Val	ctc Leu	tat Tyr	tac Tyr	tac Tyr	ctg Leu	act Thr	aag Lys	aag Lys	aat Asn	gag Glu	aac Asn	tat Tyr	aag Lys	agc Ser	1440 480
ctg Leu	gtg Val	aga Arg	cgg Arg	agc Ser	tat Tyr	cgg Arg	cgc Arg	cgc Arg	ggc Gly	aag Lys	agc Ser	cag Gln	cag Gln	caa Gln	caa Gln	1488 495
cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	cag Gln	ccc Pro	atg Met	ccc Pro	1536 510
cgc Arg	agc Ser	agc Ser	cag Gln	gag Glu	gag Glu	aaa Lys	gat Asp	gag Glu	aag Lys	gag Glu	aag Lys	gaa Glu	aag Lys	gag Glu	gcg Ala	1584 525
gag Glu	aag Lys	gag Glu	gag Glu	gag Glu	aag Lys	ccg Pro	gag Glu	gtg Val	gag Glu	aac Asn	gac Asp	aag Lys	gaa Glu	gac Asp	ctc Leu	1632 540
ctc Leu 545	aag Lys	gag Glu	aag Lys	aca Thr	gac Asp	gac Asp	acc Thr	tca Ser	ggg Gly	gag Glu	gac Asp	aac Asn	gac Asp	gag Glu	aag Lys	1680 560
gag Glu	gct Ala	gtg Val	gcc Ala	tcc Ser	aaa Lys	ggc Gly	cgc Arg	aaa Lys	act Thr	gcc Ala	aac Asn	agc Ser	cag Gln	gga Gly	aga Arg	1728 575
cgc Arg	aaa Lys	ggc Gly	cgc Arg	atc Ile	acc Thr	cgc Arg	tca Ser	atg Met	gct Ala	aat Asn	gag Glu	gcc Ala	aac Asn	agc Ser	gag Glu	1776 590
gag Glu	gcc Ala	atc Ile	acc Thr	ccc Pro	cag Gln	cag Gln	agc Ser	gcc Ala	gag Glu	ctg Leu	gcc Ala	tcc Ser	atg Met	gag Glu	ctg Leu	1824 605
aat Asn	gag Glu	agt Ser	tct Ser	cgc Arg	tgg Trp	aca Thr	gaa Glu	gaa Glu	gaa Glu	atg Met	gaa Glu	aca Thr	gcc Ala	aag Lys	aaa Lys	1872

610	615	620	
ggt ctc ctg gaa cac ggc cgc aac tgg tgc gcc atc gcc cgg atg gtg Gly Leu Leu Glu His Gly Arg Asn Trp Ser Ala Ile Ala Arg Met Val 625 630 635 640			1920
ggc tcc aag act gtg tgc cag tgt aag aac ttc tac ttc aac tac aag Gly Ser Lys Thr Val Ser Gln Cys Lys Asn Phe Tyr Phe Asn Tyr Lys 645 650 655			1968
aag agg cag aac ctc gat gag atc ttg cag cag cac aag ctg aag atg Lys Arg Gln Asn Leu Asp Glu Ile Leu Gln Gln His Lys Leu Lys Met 660 665 670			2016
gag aag gag agg aac gcg cgg agg aag aag aag aaa gcg ccg gcg gcg Glu Lys Glu Arg Asn Ala Arg Lys Lys Lys Lys Ala Pro Ala Ala 675 680 685			2064
gcc agc gag gag gct gca ttc ccg ccc gtg gtg gag gat gag gag atg Ala Ser Glu Glu Ala Ala Phe Pro Pro Val Val Glu Asp Glu Glu Met 690 695 700			2112
gag gcg tgc ggc gtg acg gga aat gag gag gag atg gtg gag gag gct Glu Ala Ser Gly Val Thr Gly Asn Glu Glu Glu Met Val Glu Glu Ala 705 710 715 720			2160
gaa gcc act gtc aac aac agc tca gac acc gag agc atc ccc tct cct Glu Ala Thr Val Asn Asn Ser Ser Asp Thr Glu Ser Ile Pro Ser Pro 725 730 735			2208
cac act gag gcc gcc aag gac aca ggg cag aat ggg ccc aag ccc cca His Thr Glu Ala Ala Lys Asp Thr Gly Gln Asn Gly Pro Lys Pro Pro 740 745 750			2256
gcc acc ctg ggc gcc gac ggg cca ccc cca ggg cca ccc acc cca cca Ala Thr Leu Gly Ala Asp Gly Pro Pro Pro Gly Pro Pro Thr Pro Pro 755 760 765			2304
ccg gag gac atc ccg gcc ccc act gag tcc acc ccg gcc tct gaa gcc Pro Glu Asp Ile Pro Ala Pro Thr Glu Ser Thr Pro Ala Ser Glu Ala 770 775 780			2352
acc tta gcc cct acg ccc cca cca gca ccc cca ttt ccc tct tca cct Thr Leu Ala Pro Thr Pro Pro Pro Ala Pro Pro Phe Pro Ser Ser Pro 785 790 795 800			2400
cct cct gtg gtc ccc aag gag gag aag gag gag gag acc gca gca gcg Pro Pro Val Val Pro Lys Glu Glu Lys Glu Glu Glu Thr Ala Ala Ala 805 810 815			2448
ccc cca gtg gag gag ggg gag gag cag aag ccc ccc gcg gct gag gag Pro Pro Val Glu Glu Gly Glu Glu Gln Lys Pro Pro Ala Ala Glu Glu 820 825 830			2496
ctg gca gtg gac aca ggg aag gcc gag gag ccc gtc aag agc gag tgc Leu Ala Val Asp Thr Gly Lys Ala Glu Glu Pro Val Lys Ser Glu Cys 835 840 845			2544
acg gag gaa gcc gag gag ggg ccg gcc aag ggc aag gac gcg gag gcc Thr Glu Glu Ala Glu Glu Gly Pro Ala Lys Gly Lys Asp Ala Glu Ala 850 855 860			2592

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted October 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gct Ala 865	gag Glu	gcc Ala	acg Thr	gcc Ala	gag Glu	agg Arg	gcg Ala	ctc Leu	aag Lys	gca Ala	gag Glu	aag Lys	aag Lys	gag Glu	ggc Gly	2640
870																
ggg Gly	agc Ser	ggc Gly	agg Arg	gcc Ala	acc Thr	aca Thr	gcc Ala	aag Lys	agc Ser	tcg Ser	ggc Gly	gcc Ala	ccc Pro	cag Gln	gac Asp	2688
885																
agc Ser	gac Asp	tcc Ser	agt Ser	gcc Ala	acc Thr	tgc Cys	agt Ser	gca Ala	gac Asp	gag Glu	gtg Val	gat Asp	gag Glu	gcc Ala	gag Glu	2736
900																
ggc Gly	ggc Gly	gac Asp	aag Lys	aac Asn	cgg Arg	ctg Leu	ctg Leu	tcc Ser	cca Pro	agg Arg	ccc Pro	agc Ser	ctc Leu	ctc Leu	acc Thr	2784
915																
ccg Pro	act Thr	ggc Gly	gac Asp	ccc Pro	cgg Arg	gcc Ala	aat Asn	gcc Ala	tca Ser	ccc Pro	cag Gln	aag Lys	cca Pro	ctg Leu	gac Asp	2832
930																
ctg Leu	aag Lys	cag Gln	ctg Leu	aag Lys	cag Gln	cga Arg	gcg Ala	gct Ala	gcc Ala	atc Ile	ccc Pro	ccc Pro	atc Ile	cag Gln	gtc Val	2880
945																
acc Thr	aaa Lys	gtc Val	cat His	gag Glu	ccc Pro	ccc Pro	cgg Arg	gag Glu	gac Asp	gca Ala	gct Ala	ccc Pro	acc Thr	aag Lys	cca Pro	2928
965																
gct Ala	ccc Pro	cca Pro	gcc Ala	cca Pro	ccg Pro	cca Pro	ccg Pro	caa Gln	aac Asn	ctg Leu	cag Gln	ccg Pro	gag Glu	agc Ser	gac Asp	2976
980																
gcc Ala	cct Pro	cag Gln	cag Gln	cct Pro	ggc Gly	agc Ser	agc Ser	ccc Pro	cgg Arg	ggc Gly	aag Lys	agc Ser	agg Arg	agc Ser	ccg Pro	3024
995																
gca Ala	ccc Pro	ccc Pro	gcc Ala	gac Asp	aag Lys	gag Glu	gca Ala	gag Glu	aag Lys	cct Pro	gtg Val	ttc Phe	ttc Phe	cca Pro	gcc Ala	3072
1010																
ttc Phe	gca Ala	gcc Ala	gag Glu	gcc Ala	cag Gln	aag Lys	ctg Leu	cct Pro	ggg Gly	gac Asp	ccc Pro	cct Pro	tgc Cys	tgg Trp	act Thr	3120
1025																
tcc Ser	ggc Gly	ctg Leu	ccc Pro	ttc Phe	ccc Pro	gtg Val	ccc Pro	ccc Pro	cgt Arg	gag Glu	gtg Val	atc Ile	aag Lys	gcc Ala	tcc Ser	3168
1040																
ccg Pro	cat His	gcc Ala	ccg Pro	gac Asp	ccc Pro	tca Ser	gcc Ala	ttc Phe	tcc Ser	tac Tyr	gct Ala	cca Pro	cct Pro	ggg Gly	cac His	3216
1060																
cca Pro	ctg Leu	ccc Pro	ctg Leu	ggc Gly	ctc Leu	cat His	gac Asp	act Thr	gcc Ala	cgg Arg	ccc Pro	gtc Val	ctg Leu	ccg Pro	cgc Arg	3264
1075																
cca Pro	ccc Pro	acc Thr	atc Ile	tcc Ser	aac Asn	ccg Pro	cct Pro	ccc Pro	ctc Leu	atc Ile	tcc Ser	tct Ser	gcc Ala	aag Lys	cac His	3312
1090																

ccc agc gtc ctc gag agg caa ata ggt gcc atc tcc caa gga atg tcg Pro Ser Val Leu Glu Arg Gln Ile Gly Ala Ile Ser Gln Gly Met Ser 1105 1110 1115 1120	3360
gtc cag ctc cac gtc ccg tac tca gag cat gcc aag gcc ccg gtg ggc Val Gln Leu His Val Pro Tyr Ser Glu His Ala Lys Ala Pro Val Gly 1125 1130 1135	3408
cct gtc acc atg ggg ctg ccc ctg ccc atg gac ccc aaa aag ctg gca Pro Val Thr Met Gly Leu Pro Leu Pro Met Asp Pro Lys Lys Leu Ala 1140 1145 1150	3456
ccc ttc agc gga gtg aag cag gag cag ctg tcc cca cgg ggc cag gct Pro Phe Ser Gly Val Lys Gln Glu Gln Leu Ser Pro Arg Gly Gln Ala 1155 1160 1165	3504
ggg cca ccg gag agc ctg ggg gtg ccc aca gcc cag gag gcg tcc gtg Gly Pro Pro Glu Ser Leu Gly Val Pro Thr Ala Gln Glu Ala Ser Val 1170 1175 1180	3552
ctg aga ggg aca gct ctg ggc tca gtt ccg ggc gga agc atc acc aaa Leu Arg Gly Thr Ala Leu Gly Ser Val Pro Gly Gly Ser Ile Thr Lys 1185 1190 1195 1200	3600
ggc att ccc agc aca cgg gtg ccc tcg gac agc gcc atc aca tac cgc Gly Ile Pro Ser Thr Arg Val Pro Ser Asp Ser Ala Ile Thr Tyr Arg 1205 1210 1215	3648
ggc tcc atc acc cac ggc acg cca gct gac gtc ctg tac aag ggc acc Gly Ser Ile Thr His Gly Thr Pro Ala Asp Val Leu Tyr Lys Gly Thr 1220 1225 1230	3696
atc acc agg atc atc ggc gag gac agc ccg agt cgc ttg gac cgc ggc Ile Thr Arg Ile Ile Gly Glu Asp Ser Pro Ser Arg Leu Asp Arg Gly 1235 1240 1245	3744
cgg gag gac agc ctg ccc aag ggc cac gtc atc tac gaa ggc aag aag Arg Glu Asp Ser Leu Pro Lys Gly His Val Ile Tyr Glu Gly Lys Lys 1250 1255 1260	3792
ggc cac gtc ttg tcc tat gag ggt ggc atg tct gtg acc cag tgc tcc Gly His Val Leu Ser Tyr Glu Gly Gly Met Ser Val Thr Gln Cys Ser 1265 1270 1275 1280	3840
aag gag gac ggc aga agc agc tca gga ccc ccc cat gag acg gcc gcc Lys Glu Asp Gly Arg Ser Ser Ser Gly Pro Pro His Glu Thr Ala Ala 1285 1290 1295	3888
ccc aag cgc acc tat gac atg atg gag ggc cgc gtg ggc aga gcc atc Pro Lys Arg Thr Tyr Asp Met Met Glu Gly Arg Val Gly Arg Ala Ile 1300 1305 1310	3936
tcc tca gcc agc atc gaa ggt ctc atg ggc cgt gcc atc ccg ccg gag Ser Ser Ala Ser Ile Glu Gly Leu Met Gly Arg Ala Ile Pro Pro Glu 1315 1320 1325	3984
cga cac agc ccc cac cac ctc aaa gag cag cac cac atc cgc ggg tcc Arg His Ser Pro His His Leu Lys Glu Gln His His Ile Arg Gly Ser 1330 1335 1340	4032
atc aca caa ggg atc cct ccg tcc tac gtg gag gca cag gag gac tac	4080

Ile Thr Gln Gly Ile Pro Arg Ser Tyr Val Glu Ala Gln Glu Asp Tyr			
1345	1350	1355	1360
ctg cgt cgg gag gcc aag ctc cta aag cgg gag ggc acg cct ccg ccc	4128		
Leu Arg Arg Glu Ala Lys Leu Leu Lys Arg Glu Gly Thr Pro Pro Pro			
1365	1370	1375	
cca ccg ccc tca cgg gac ctg acc gag gcc tac aag acg cag gcc ctg	4176		
Pro Pro Pro Ser Arg Asp Leu Thr Glu Ala Tyr Lys Thr Gln Ala Leu			
1380	1385	1390	
ggc ccc ctg aag ctg aag ccg gcc cat gag ggc ctg gtg gcc acg gtg	4224		
Gly Pro Leu Lys Leu Lys Pro Ala His Glu Gly Leu Val Ala Thr Val			
1395	1400	1405	
aag gag gcg ggc cgc tcc atc cat gag atc ccg cgc gag gag ctg cgg	4272		
Lys Glu Ala Gly Arg Ser Ile His Glu Ile Pro Arg Glu Glu Leu Arg			
1410	1415	1420	
cac acg ccc gag ctg ccc ctg gcc ccg cgg ccg ctc aag gag ggc tcc	4320		
His Thr Pro Glu Leu Pro Leu Ala Pro Arg Pro Leu Lys Glu Gly Ser			
1425	1430	1435	1440
atc acg cag ggc acc ccg ctc aag tac gac acc ggc gcg tcc acc act	4368		
Ile Thr Gln Gly Thr Pro Leu Lys Tyr Asp Thr Gly Ala Ser Thr Thr			
1445	1450	1455	
ggc tcc aaa aag cac gac gta cgc tcc ctc atc ggc agc ccc ggc cgg	4416		
Gly Ser Lys Lys His Asp Val Arg Ser Leu Ile Gly Ser Pro Gly Arg			
1460	1465	1470	
acg ttc cca ccc gtg cac ccg ctg gat gtg atg gcc gac gcc cgg gca	4464		
Thr Phe Pro Pro Val His Pro Leu Asp Val Met Ala Asp Ala Arg Ala			
1475	1480	1485	
ctg gaa cgt gcc tgc tac gag gag agc ctg aag agc cgg cca ggg acc	4512		
Leu Glu Arg Ala Cys Tyr Glu Glu Ser Leu Lys Ser Arg Pro Gly Thr			
1490	1495	1500	
gcc agc agc tcg ggg ggc tcc att gcg cgc ggc gcc ccg gtc att gtg	4560		
Ala Ser Ser Ser Gly Gly Ser Ile Ala Arg Gly Ala Pro Val Ile Val			
1505	1510	1515	1520
cct gag ctg ggt aag ccg cgg cag agc ccc ctg acc tat gag gac cac	4608		
Pro Glu Leu Gly Lys Pro Arg Gln Ser Pro Leu Thr Tyr Glu Asp His			
1525	1530	1535	
ggg gca ccc ttt gcc ggc cac ctc cca cga ggt tcg ccc gtg acc atg	4656		
Gly Ala Pro Phe Ala Gly His Leu Pro Arg Gly Ser Pro Val Thr Met			
1540	1545	1550	
cgg gag ccc acg ccg cgc ctg cag gag ggc agc ctt tcg tcc agc aag	4704		
Arg Glu Pro Thr Pro Arg Leu Gln Glu Gly Ser Leu Ser Ser Ser Lys			
1555	1560	1565	
gca tcc cag gac cga aag ctg acg tcg acg cct cgt gag atc gcc aag	4752		
Ala Ser Gln Asp Arg Lys Leu Thr Ser Thr Pro Arg Glu Ile Ala Lys			
1570	1575	1580	
tcc ccg cac agc acc gtg ccc gag cac cac cca cac ccc atc tcg ccc	4800		
Ser Pro His Ser Thr Val Pro Glu His His Pro His Pro Ile Ser Pro			

1585	1590										1595										1600																			
tat gag cac ctg ctt cgg ggc gtg agt ggc gtg gac ctg tat cgc agc	Tyr Glu His Leu Leu Arg Gly Val Ser Gly Val Asp Leu Tyr Arg Ser																														4848									
1605										1610										1615																				
cac atc ccc ctg gcc ttc gac ccc acc tcc ata ccc cgc ggc atc cct	His Ile Pro Leu Ala Phe Asp Pro Thr Ser Ile Pro Arg Gly Ile Pro																														4896									
1620										1625										1630																				
ctg gac gca gcc gct gcc tac tac ctg ccc cga cac ctg gcc ccc aac	Leu Asp Ala Ala Ala Ala Tyr Tyr Leu Pro Arg His Leu Ala Pro Asn																														4944									
1635										1640										1645																				
ccc acc tac ccg cac ctg tac cca ccc tac ctc atc cgc ggc tac ccc	Pro Thr Tyr Pro His Leu Tyr Pro Pro Tyr Leu Ile Arg Gly Tyr Pro																														4992									
1650										1655										1660																				
gac acg gcg gcg ctg gag aac cgg cag acc atc atc aat gac tac atc	Asp Thr Ala Ala Leu Glu Asn Arg Gln Thr Ile Ile Asn Asp Tyr Ile																														5040									
1665										1670										1675										1680										
acc tcg cag cag atg cac cac aac acg gcc acc gcc atg gcc cag cga	Thr Ser Gln Gln Met His His Asn Thr Ala Thr Ala Met Ala Gln Arg																														5088									
1685										1690										1695																				
gct gat atg ctg agg ggc ctc tcg ccc cgc gag tcc tcg ctg gca ctc	Ala Asp Met Leu Arg Gly Leu Ser Pro Arg Glu Ser Ser Leu Ala Leu																														5136									
1700										1705										1710																				
aac tac gct gcg ggt ccc cga ggc atc atc gac ctg tcc caa gtg cca	Asn Tyr Ala Ala Gly Pro Arg Gly Ile Ile Asp Leu Ser Gln Val Pro																														5184									
1715										1720										1725																				
cac ctg cct gtg ctc gtg ccc ccg aca cca ggc acc cca gcc acc gcc	His Leu Pro Val Leu Val Pro Pro Thr Pro Gly Thr Pro Ala Thr Ala																														5232									
1730										1735										1740																				
atg gac cgc ctt gcc tac ctc ccc acc gcg ccc cag ccc ttc agc agc	Met Asp Arg Leu Ala Tyr Leu Pro Thr Ala Pro Gln Pro Phe Ser Ser																														5280									
1745										1750										1755										1760										
cgc cac agc agc tcc cca ctc tcc cca gga ggt cca aca cac ttg aca	Arg His Ser Ser Ser Pro Leu Ser Pro Gly Gly Pro Thr His Leu Thr																														5328									
1765										1770										1775																				
aaa cca acc acc acg tcc tcg tcc gag cgg gag cga gac cgg gat cga	Lys Pro Thr Thr Thr Ser Ser Ser Ser Glu Arg Glu Arg Asp Arg Asp Arg																														5376									
1780										1785										1790																				
gag cgg gac cgg gat cgg gag cgg gaa aag tcc atc ctc acg tcc acc	Glu Arg Asp Arg Asp Arg Glu Arg Glu Lys Ser Ile Leu Thr Ser Thr																														5424									
1795										1800										1805																				
acg acg gtg gag cac gca ccc atc tgg aga cct ggt aca gag cag agc	Thr Thr Val Glu His Ala Pro Ile Trp Arg Pro Gly Thr Glu Gln Ser																														5472									
1810										1815										1820																				
agc ggc agc agc ggc agc agc ggc ggg ggt ggg ggc agc agc agc cgc	Ser Gly Ser Ser Gly Ser Ser Gly Gly Gly Gly Gly Ser Ser Ser Arg																														5520									
1825										1830										1835										1840										

cca ggc ccc gtg aag ctt ggc ggg gag gcc gcc cac ctc cca cac ctg Pro Gly Pro Val Lys Leu Gly Gly Glu Ala Ala His Leu Pro His Leu	6288
2085 2090 2095	
cgg ccg ctg cct gag agc cag ccc tcg tcc agc ccg ctg ctc cag acc Arg Pro Leu Pro Glu Ser Gln Pro Ser Ser Ser Pro Leu Leu Gln Thr	6336
2100 2105 2110	
gcc cca ggg gtc aaa ggt cac cag cgg gtg gtc acc ctg gcc cag cac Ala Pro Gly Val Lys Gly His Gln Arg Val Val Thr Leu Ala Gln His	6384
2115 2120 2125	
atc agt gag gtc atc aca cag gac tac acc cgg cac cac cca cag cag Ile Ser Glu Val Ile Thr Gln Asp Tyr Thr Arg His His Pro Gln Gln	6432
2130 2135 2140	
ctc agc gca ccc ctg ccc gcc ccc ctc tac tcc ttc cct ggg gcc agc Leu Ser Ala Pro Leu Pro Ala Pro Leu Tyr Ser Phe Pro Gly Ala Ser	6480
2145 2150 2155 2160	
tgc ccc gtc ctg gac ctc cgc cgc cca ccc agt gac ctc tac ctc ccg Cys Pro Val Leu Asp Leu Arg Arg Pro Pro Ser Asp Leu Tyr Leu Pro	6528
2165 2170 2175	
ccc ccg gac cat ggt gcc ccg gcc cgt ggc tcc ccc cac agc gaa ggg Pro Pro Asp His Gly Ala Pro Ala Arg Gly Ser Pro His Ser Glu Gly	6576
2180 2185 2190	
ggc aag agg tct cca gag cca aac aag acg tcg gtc ttg ggt ggt ggt Gly Lys Arg Ser Pro Glu Pro Asn Lys Thr Ser Val Leu Gly Gly Gly	6624
2195 2200 2205	
gag gac ggt att gaa cct gtg tcc cca ccg gag ggc atg acg gag cca Glu Asp Gly Ile Glu Pro Val Ser Pro Pro Glu Gly Met Thr Glu Pro	6672
2210 2215 2220	
ggg cac tcc cgg agt gct gtg tac ccg ctg ctg tac cgg gat ggg gaa Gly His Ser Arg Ser Ala Val Tyr Pro Leu Leu Tyr Arg Asp Gly Glu	6720
2225 2230 2235 2240	
cag acg gag ccc agc agg atg ggc tcc aag tct cca ggc aac acc agc Gln Thr Glu Pro Ser Arg Met Gly Ser Lys Ser Pro Gly Asn Thr Ser	6768
2245 2250 2255	
cag ccg cca gcc ttc ttc agc aag ctg acc gag agc aac tcc gcc atg Gln Pro Pro Ala Phe Phe Ser Lys Leu Thr Glu Ser Asn Ser Ala Met	6816
2260 2265 2270	
gtc aag tcc aag aag caa gag atc aac aag aag ctg aac acc cac aac Val Lys Ser Lys Lys Gln Glu Ile Asn Lys Lys Leu Asn Thr His Asn	6864
2275 2280 2285	
cgg aat gag cct gaa tac aat atc agc cag cct ggg acg gag atc ttc Arg Asn Glu Pro Glu Tyr Asn Ile Ser Gln Pro Gly Thr Glu Ile Phe	6912
2290 2295 2300	
aat atg ccc gcc atc acc gga aca ggc ctt atg acc tat aga agc cag Asn Met Pro Ala Ile Thr Gly Thr Gly Leu Met Thr Tyr Arg Ser Gln	6960
2305 2310 2315 2320	
gcg gtg cag gaa cat gcc agc acc aac atg ggg ctg gag gcc ata att	7008

Ala Val Gln Glu His Ala Ser Thr Asn Met Gly Leu Glu Ala Ile Ile	
2325 2330 2335	
aga aag gca ctc atg ggt aaa tat gac cag tgg gaa gag tcc ccg ccg	7056
Arg Lys Ala Leu Met Gly Lys Tyr Asp Gln Trp Glu Glu Ser Pro Pro	
2340 2345 2350	
ctc agc gcc aat gct ttt aac cct ctg aat gcc agt gcc agc ctg ccc	7104
Leu Ser Ala Asn Ala Phe Asn Pro Leu Asn Ala Ser Ala Ser Leu Pro	
2355 2360 2365	
gct gct atg ccc ata acc gct gct gac gga cgg agt gac cac aca ctc	7152
Ala Ala Met Pro Ile Thr Ala Ala Asp Gly Arg Ser Asp His Thr Leu	
2370 2375 2380	
acc tcg cca ggt ggc ggc ggg aag gcc aag gtc tct ggc aga ccc agc	7200
Thr Ser Pro Gly Gly Gly Lys Ala Lys Val Ser Gly Arg Pro Ser	
2385 2390 2395 2400	
agc cga aaa gcc aag tcc ccg gcc ccg ggc ctg gca tct ggg gac cgg	7248
Ser Arg Lys Ala Lys Ser Pro Ala Pro Gly Leu Ala Ser Gly Asp Arg	
2405 2410 2415	
cca ccc tct gtc tcc tca gtg cac tcg gag gga gac tgc aac cgc cgg	7296
Pro Pro Ser Val Ser Ser Val His Ser Glu Gly Asp Cys Asn Arg Arg	
2420 2425 2430	
acg ccg ctc acc aac cgc gtg tgg gag gac agg ccc tcg tcc gca ggt	7344
Thr Pro Leu Thr Asn Arg Val Trp Glu Asp Arg Pro Ser Ser Ala Gly	
2435 2440 2445	
tcc acg cca ttc ccc tac aac ccc ctg atc atg cgg ctg cag gcg ggt	7392
Ser Thr Pro Phe Pro Tyr Asn Pro Leu Ile Met Arg Leu Gln Ala Gly	
2450 2455 2460	
gtc atg gct tcc cca ccc cca ccg ggc ctc ccc gcg ggc agc ggg ccc	7440
Val Met Ala Ser Pro Pro Pro Pro Gly Leu Pro Ala Gly Ser Gly Pro	
2465 2470 2475 2480	
ctc gct ggc ccc cac cac gcc tgg gac gag gag ccc aag cca ctg ctc	7488
Leu Ala Gly Pro His His Ala Trp Asp Glu Glu Pro Lys Pro Leu Leu	
2485 2490 2495	
tgc tcg cag tac gag aca ctc tcc gac agc gag	7521
Cys Ser Gln Tyr Glu Thr Leu Ser Asp Ser Glu	
2500 2505	

<210> 4
 <211> 8544
 <212> DNA
 <213> Mus musculus

<220>
 <221> CDS
 <222> (160)..(7545)

<400> 4
 ctcgagcccg atccgccgta gcccgccgcc agcgcccggt gccgccgccg ggaggcacct 60
 gtgacgaggt cacctgccag cagatgaccg agaccagccc ttagtcctag gtgtggtcaa 120

gagtgctcttg gctccaaagc ctacctggac cctaccacc atg tca gga tcc aca															174				
Met Ser Gly Ser Thr																			
1 5																			
cag cct gtg gca cag aca tgg cgg gct gct gag ccc cgc tac cca ccc	222																		
Gln Pro Val Ala Gln Thr Trp Arg Ala Ala Glu Pro Arg Tyr Pro Pro																			
10					15					20									
cat ggc atc tcc tac ccg gtg cag ata gcc cgg tcc cac acg gac gtg	270																		
His Gly Ile Ser Tyr Pro Val Gln Ile Ala Arg Ser His Thr Asp Val																			
25					30					35									
ggg ctg ctt gag tac caa cac cac ccc cgt gac tac acc tca cac ctg	318																		
Gly Leu Leu Glu Tyr Gln His His Pro Arg Asp Tyr Thr Ser His Leu																			
40					45					50									
tca ccc ggt tcc atc atc cag cca cag agg agg cgg ccc tca ctg ctg	366																		
Ser Pro Gly Ser Ile Ile Gln Pro Gln Arg Arg Arg Pro Ser Leu Leu																			
55					60					65									
tca gag ttc cag cct ggg agt gaa cgg tct cag gag ctc cac ctg cgc	414																		
Ser Glu Phe Gln Pro Gly Ser Glu Arg Ser Gln Glu Leu His Leu Arg																			
70					75					80					85				
cct gag tcc cgc acg ttc ctg cct gag ctg ggc aag ccc gac ata gaa	462																		
Pro Glu Ser Arg Thr Phe Leu Pro Glu Leu Gly Lys Pro Asp Ile Glu																			
90					95					100									
ttc acc gag agc aag cgc ccc cgc ctg gag cta cta ccc gat acc ctg	510																		
Phe Thr Glu Ser Lys Arg Pro Arg Leu Glu Leu Leu Pro Asp Thr Leu																			
105					110					115									
ctg cgc cca tca ccc ctg ctg gcc act ggg cag ccg agt ggg tct gaa	558																		
Leu Arg Pro Ser Pro Leu Leu Ala Thr Gly Gln Pro Ser Gly Ser Glu																			
120					125					130									
gac ctt acc aag gac cgt agc ctg gca ggc aag ctg gag cct gtg tca	606																		
Asp Leu Thr Lys Asp Arg Ser Leu Ala Gly Lys Leu Glu Pro Val Ser																			
135					140					145									
cct ccc agt ccc ccg cac gct gac cct gag cta gag ctg gcg cca tct	654																		
Pro Pro Ser Pro Pro His Ala Asp Pro Glu Leu Glu Leu Ala Pro Ser																			
150					155					160					165				
cga ctg tcc aag gag gag ctg atc cag aac aga ttg gac cgc gtg gac	702																		
Arg Leu Ser Lys Glu Glu Leu Ile Gln Asn Arg Leu Asp Arg Val Asp																			
170					175					180									
cgt gag atc acc atg gta gag cag cag atc tcc aag ctg aag aag aag	750																		
Arg Glu Ile Thr Met Val Glu Gln Gln Ile Ser Lys Leu Lys Lys Lys																			
185					190					195									
cag caa cag ttg gag gag gag gcc gcc aag ccg ccc gaa ccc gag aag	798																		
Gln Gln Gln Leu Glu Glu Glu Ala Ala Lys Pro Pro Glu Pro Glu Lys																			
200					205					210									
cct gtg tcg cca cca ccc ata gaa tca aag cac cga agc ctg gtc cag	846																		
Pro Val Ser Pro Pro Pro Ile Glu Ser Lys His Arg Ser Leu Val Gln																			
215					220					225									
atc atc tac gat gag aac cgg aag aaa gcc gaa gcc gca cac cgg atc	894																		

Ile	Ile	Tyr	Asp	Glu	Asn	Arg	Lys	Lys	Ala	Glu	Ala	Ala	His	Arg	Ile	
230					235					240					245	
cta	gaa	ggc	ctg	ggg	ccc	cag	gtg	gag	ctg	cct	ctg	tac	aac	cag	ccg	942
Leu	Glu	Gly	Leu	Gly	Pro	Gln	Val	Glu	Leu	Pro	Leu	Tyr	Asn	Gln	Pro	
				250					255					260		
tct	gac	aca	cgc	cag	tac	cat	gaa	aac	atc	aaa	ata	aac	cag	gcg	atg	990
Ser	Asp	Thr	Arg	Gln	Tyr	His	Glu	Asn	Ile	Lys	Ile	Asn	Gln	Ala	Met	
			265					270					275			
cgg	aag	aag	ctg	atc	ttg	tac	ttt	aag	cgg	agg	aac	cac	gcg	cgc	aag	1038
Arg	Lys	Lys	Leu	Ile	Leu	Tyr	Phe	Lys	Arg	Arg	Asn	His	Ala	Arg	Lys	
		280					285					290				
cag	tgg	gaa	cag	cgc	ttc	tgc	cag	cgc	tat	gac	cag	ctc	atg	gag	gcg	1086
Gln	Trp	Glu	Gln	Arg	Phe	Cys	Gln	Arg	Tyr	Asp	Gln	Leu	Met	Glu	Ala	
	295					300					305					
tgg	gag	aag	aag	gta	gag	cgc	ata	gag	aac	aat	ccg	cga	agg	agg	gcc	1134
Trp	Glu	Lys	Lys	Val	Glu	Arg	Ile	Glu	Asn	Asn	Pro	Arg	Arg	Arg	Ala	
310					315					320					325	
aag	gag	agc	aag	gtg	agg	gag	tac	tac	gag	aaa	cag	ttc	ccg	gag	atc	1182
Lys	Glu	Ser	Lys	Val	Arg	Glu	Tyr	Tyr	Glu	Lys	Gln	Phe	Pro	Glu	Ile	
				330					335					340		
cgc	aag	cag	cgg	gag	ctg	cag	gag	cgc	atg	cag	agc	agg	gtg	ggc	cag	1230
Arg	Lys	Gln	Arg	Glu	Leu	Gln	Glu	Arg	Met	Gln	Ser	Arg	Val	Gly	Gln	
			345					350					355			
cgt	ggc	agt	ggg	ctc	tcc	atg	tgc	gct	gcc	cgc	agt	gag	cat	gag	gtt	1278
Arg	Gly	Ser	Gly	Leu	Ser	Met	Ser	Ala	Ala	Arg	Ser	Glu	His	Glu	Val	
		360					365					370				
tct	gag	atc	att	gat	ggc	ttg	tct	gag	cag	gag	aac	ctg	gag	aag	cag	1326
Ser	Glu	Ile	Ile	Asp	Gly	Leu	Ser	Glu	Gln	Glu	Asn	Leu	Glu	Lys	Gln	
	375					380					385					
atg	cgc	cag	ctg	gcc	gtg	atc	cgc	cat	gtt	gta	cga	cgc	gac	cag	cag	1374
Met	Arg	Gln	Leu	Ala	Val	Ile	Arg	His	Val	Val	Arg	Arg	Asp	Gln	Gln	
	390				395					400					405	
agg	atc	aag	ttc	atc	aac	atg	aat	gga	ctc	atg	gat	gac	ccc	atg	aag	1422
Arg	Ile	Lys	Phe	Ile	Asn	Met	Asn	Gly	Leu	Met	Asp	Asp	Pro	Met	Lys	
				410				415						420		
gtc	tac	aag	gac	cgt	cag	gtt	acc	aac	atg	tgg	agc	gag	cag	gag	agg	1470
Val	Tyr	Lys	Asp	Arg	Gln	Val	Thr	Asn	Met	Trp	Ser	Glu	Gln	Glu	Arg	
			425					430					435			
gac	acc	ttc	cgt	gag	aag	ttt	atg	cag	cac	cct	aag	aac	ttt	ggc	ctg	1518
Asp	Thr	Phe	Arg	Glu	Lys	Phe	Met	Gln	His	Pro	Lys	Asn	Phe	Gly	Leu	
		440					445					450				
att	gcc	tca	ttc	ctg	gag	aga	aag	acg	gtc	gct	gag	tgt	gtc	ctc	tat	1566
Ile	Ala	Ser	Phe	Leu	Glu	Arg	Lys	Thr	Val	Ala	Glu	Cys	Val	Leu	Tyr	
	455					460					465					
tac	tac	ctg	acc	aag	aag	aat	gaa	aat	tac	aag	agc	ttg	gtg	agg	cgg	1614
Tyr	Tyr	Leu	Thr	Lys	Lys	Asn	Glu	Asn	Tyr	Lys	Ser	Leu	Val	Arg	Arg	

470	475	480	485	
agc tat cgg cgc cgt ggc aag agc cag cag cag cag cag cag caa caa				1662
Ser Tyr Arg Arg Arg Gly Lys Ser Gln Gln Gln Gln Gln Gln Gln Gln	490	495	500	
cag cag cag cag cag cag atg gca cgg agc agc cag gag gag aag gag				1710
Gln Gln Gln Gln Gln Gln Met Ala Arg Ser Ser Gln Glu Glu Lys Glu	505	510	515	
gag aag gag aag gag aag gag gcc gac aag gag gaa gag aag cag gat				1758
Glu Lys Glu Lys Glu Lys Glu Ala Asp Lys Glu Glu Glu Lys Gln Asp	520	525	530	
gcg gag aac gag aag gaa gaa ctc agc aag gag aag aca gac gac act				1806
Ala Glu Asn Glu Lys Glu Glu Leu Ser Lys Glu Lys Thr Asp Asp Thr	535	540	545	
tct gcc gag gac aac gat gag aaa gag gcc gtg gcc tcc aaa gcc cgc				1854
Ser Gly Glu Asp Asn Asp Glu Lys Glu Ala Val Ala Ser Lys Gly Arg	550	555	560	565
aaa act gcc aac agc caa gcc cgc cgc aaa gcc cgt atc acg cgc tcc				1902
Lys Thr Ala Asn Ser Gln Gly Arg Arg Lys Gly Arg Ile Thr Arg Ser	570	575	580	
atg gcc aac gag gcc aac cat gag gag aca gcc acc cca cag caa agt				1950
Met Ala Asn Glu Ala Asn His Glu Glu Thr Ala Thr Pro Gln Gln Ser	585	590	595	
tca gag ctg gct tcc atg gag atg aac gag agt tct cgc tgg act gag				1998
Ser Glu Leu Ala Ser Met Glu Met Asn Glu Ser Ser Arg Trp Thr Glu	600	605	610	
gaa gag atg gag aca gca aag aaa gcc ctc ctg gaa cat ggg agg aac				2046
Glu Glu Met Glu Thr Ala Lys Lys Gly Leu Leu Glu His Gly Arg Asn	615	620	625	
tgg tca gcc att gcc cgc atg gtg gcc tcc aag acc gtg tcc cag tgt				2094
Trp Ser Ala Ile Ala Arg Met Val Gly Ser Lys Thr Val Ser Gln Cys	630	635	640	645
aag aac ttc tac ttc aac tac aag aag agg cag aac ctg gac gaa atc				2142
Lys Asn Phe Tyr Phe Asn Tyr Lys Lys Arg Gln Asn Leu Asp Glu Ile	650	655	660	
ctt cag cag cac aag cta aag atg gag aag gag agg aac gct cgg agg				2190
Leu Gln Gln His Lys Leu Lys Met Glu Lys Glu Arg Asn Ala Arg Arg	665	670	675	
aag aag aag aag acc cca gct gcg gcg agc gag gag aca gcc ttc cca				2238
Lys Lys Lys Lys Thr Pro Ala Ala Ala Ser Glu Glu Thr Ala Phe Pro	680	685	690	
cct gcc gct gag gac gaa gag atg gaa gca tca gcc gca agt gcc aat				2286
Pro Ala Ala Glu Asp Glu Glu Met Glu Ala Ser Gly Ala Ser Ala Asn	695	700	705	
gag gaa gag ctg gcg gag gag gca gaa gcc tca cag gcc tct ggg aat				2334
Glu Glu Glu Leu Ala Glu Glu Ala Glu Ala Ser Gln Ala Ser Gly Asn	710	715	720	725

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gag gtt ccc aga gtt ggg gag tgc agt ggc cca gct gct gtc aac aac	2382
Glu Val Pro Arg Val Gly Glu Cys Ser Gly Pro Ala Ala Val Asn Asn	
730 735 740	
agc tct gat act gag agt gtc cca tcc ccg cgt tca gaa gcc acg aag	2430
Ser Ser Asp Thr Glu Ser Val Pro Ser Pro Arg Ser Glu Ala Thr Lys	
745 750 755	
gac act ggg cct aaa ccc act ggc act gaa gca ttg ccc gct gcc acc	2478
Asp Thr Gly Pro Lys Pro Thr Gly Thr Glu Ala Leu Pro Ala Ala Thr	
760 765 770	
cag cca cct gtt cct cct cca gaa gaa ccg gca gca gcc cct gct gag	2526
Gln Pro Pro Val Pro Pro Pro Glu Glu Pro Ala Ala Ala Pro Ala Glu	
775 780 785	
ccc tcc cca gtc cct gat gcc agt ggc cca cca tcc cca gag cct tcc	2574
Pro Ser Pro Val Pro Asp Ala Ser Gly Pro Pro Ser Pro Glu Pro Ser	
790 795 800 805	
cca tca cct gcc gca ccc ccg gct act gtg gac aag gat gaa caa gaa	2622
Pro Ser Pro Ala Ala Pro Pro Ala Thr Val Asp Lys Asp Glu Gln Glu	
810 815 820	
gcc ccg gct gct cca gct ccc cag aca gaa gat gcc aag gag cag aag	2670
Ala Pro Ala Ala Pro Ala Pro Gln Thr Glu Asp Ala Lys Glu Gln Lys	
825 830 835	
tct gag gcc gag gag atc gat gtg gga aag cca gag gag ccc gag gcc	2718
Ser Glu Ala Glu Glu Ile Asp Val Gly Lys Pro Glu Glu Pro Glu Ala	
840 845 850	
tct gag gag ccc ccg gag agt gta aag agt gac cac aag gag gag acc	2766
Ser Glu Glu Pro Pro Glu Ser Val Lys Ser Asp His Lys Glu Glu Thr	
855 860 865	
gag gaa gag cct gaa gac aaa gcc aag ggc aca gag gcc att gaa act	2814
Glu Glu Glu Pro Glu Asp Lys Ala Lys Gly Thr Glu Ala Ile Glu Thr	
870 875 880 885	
gtg tct gag gca cca ctt aag gtg gag gag gct ggt agc aag gca gct	2862
Val Ser Glu Ala Pro Leu Lys Val Glu Glu Ala Gly Ser Lys Ala Ala	
890 895 900	
gtg acc aag ggt tcc agc tca ggt gcc acc cag gac agt gac tcc agt	2910
Val Thr Lys Gly Ser Ser Ser Gly Ala Thr Gln Asp Ser Asp Ser Ser	
905 910 915	
gcc acc tgc agt gcc gat gag gtg gac gaa ccc gaa gga ggt gac aag	2958
Ala Thr Cys Ser Ala Asp Glu Val Asp Glu Pro Glu Gly Gly Asp Lys	
920 925 930	
ggc agg ctg ctg tca cca agg ccc agc ctc ctc acc ccg gct gga gat	3006
Gly Arg Leu Leu Ser Pro Arg Pro Ser Leu Leu Thr Pro Ala Gly Asp	
935 940 945	
ccc cgg gcc agt acc tcg ccc cag aag ccg ctg gac ctg aag cag ctg	3054
Pro Arg Ala Ser Thr Ser Pro Gln Lys Pro Leu Asp Leu Lys Gln Leu	
950 955 960 965	

2025-10-16 16:00

aag Lys	cag Gln	cga Arg	gca Ala	gcc Ala	gcc Ala	atc Ile	ccc Pro	ccc Pro	cct Ile	atc Val	gtc Thr	acc Lys	aag Val	gtc His	cat Glu	gag	3102
970 975 980																	
ccc Pro	ccc Pro	cgg Arg	gag Glu	gac Asp	aca Thr	gta Val	ccc Pro	cca Pro	aag Lys	cca Pro	gtt Val	ccc Pro	cct Pro	gtg Val	cct Pro		3150
985 990 995																	
cca Pro	ccc Pro	acg Thr	cag Gln	cac His	cta Leu	cag Gln	cca Pro	gag Glu	ggg Gly	gac Asp	gtg Val	tct Ser	cag Gln	cag Gln	tcg Ser		3198
1000 1005 1010																	
gga Gly	gga Gly	agt Ser	cca Pro	cgt Arg	ggc Gly	aag Lys	tcc Ser	cgc Arg	agc Ser	cca Pro	gtg Val	cct Pro	cct Pro	gcc Ala	gag Glu		3246
1015 1020 1025																	
aaa Lys	gag Glu	gca Ala	gag Glu	aaa Lys	ccc Pro	gca Ala	ttc Phe	ttt Phe	ccg Pro	gct Ala	ttc Phe	cca Pro	act Thr	gag Glu	ggc Gly		3294
1030 1035 1040 1045																	
caa Gln	agc Ser	tac Tyr	cga Arg	ctg Leu	agc Ser	ccc Pro	cac His	gct Ala	ggg Gly	cat His	cgg Arg	ctg Leu	cct Pro	tcc Ser	cat His		3342
1050 1055 1060																	
cct Pro	cca Pro	cgg Arg	gag Glu	gtg Val	atc Ile	aag Lys	act Thr	tcc Ser	aca Thr	cgc Arg	gct Ala	gac Asp	cct Pro	ctc Leu	ttc Phe		3390
1065 1070 1075																	
tcc Ser	tac Tyr	aca Thr	ccc Pro	ccc Pro	ggg Gly	cac His	ccg Pro	ctg Leu	cct Pro	ctg Leu	ggc Gly	ctc Leu	cac His	gat Asp	agt Ser		3438
1080 1085 1090																	
gcc Ala	cgg Arg	ccc Pro	gtc Val	ctg Leu	cca Pro	cgt Arg	ccc Pro	ccc Pro	atc Ile	tct Ser	aac Asn	ccc Pro	cca Pro	ccc Pro	ctc Leu		3486
1095 1100 1105																	
atc Ile	tcc Ser	tct Ser	gcc Ala	aag Lys	cat His	ccc Pro	ggc Gly	gta Val	ctt Leu	gag Glu	agg Arg	cag Gln	ctg Leu	ggg Gly	gcc Ala		3534
1110 1115 1120 1125																	
atc Ile	tcc Ser	cag Gln	ggg Gly	atg Met	tca Ser	gtc Val	cag Gln	ctt Leu	cgt Arg	gtg Val	cct Pro	cac His	tca Ser	gag Glu	cat His		3582
1130 1135 1140																	
gcc Ala	aag Lys	ccc Pro	atg Met	ggc Gly	cct Pro	ctc Leu	acc Thr	atg Met	gag Glu	ctg Leu	ccc Pro	ctt Leu	gcc Ala	gtg Val	gac Asp		3630
1145 1150 1155																	
cct Pro	aag Lys	aag Lys	ctg Leu	ggg Gly	aca Thr	gca Ala	ctg Leu	gct Ala	ccg Pro	cca Pro	cca Pro	gtg Val	gaa Glu	gca Ala	tca Ser		3678
1160 1165 1170																	
cca Pro	agg Arg	gcc Ala	tcc Ser	cag Gln	tac Tyr	ccg Pro	ggc Gly	tgc Cys	aga Arg	cgg Arg	ccc Pro	cag Gln	cta Leu	cag Gln	agg Arg		3726
1175 1180 1185																	
ctc Leu	tat Tyr	cac His	cca Pro	cgc Arg	acg Thr	ccc Pro	gca Ala	gac Asp	gtc Val	ctc Leu	tac Tyr	aag Lys	ggg Gly	acc Thr	atc Ile		3774
1190 1195 1200 1205																	
agc Arg	agg Arg	atc Arg	gtc Arg	ggg Arg	gag Gly	gac Gly	agc Arg	cca Arg	agt Arg	cgc Arg	ctt Arg	gac Arg	cgg Arg	gca Arg	cga Arg		3822

gag gac acc ctg ccc aag ggc cat gtc atc tat gag ggc aag aaa ggc	3870
Glu Asp Thr Leu Pro Lys Gly His Val Ile Tyr Glu Gly Lys Lys Gly	
1225 1230 1235	
cac gtc cta tcc tat gaa ggt ggt atg tcc gtg tca cag tgc tct aag	3918
His Val Leu Ser Tyr Glu Gly Gly Met Ser Val Ser Gln Cys Ser Lys	
1240 1245 1250	
gag gat gga agg agc agc tcg ggc cca ccc cat gag act gcc gcc cct	3966
Glu Asp Gly Arg Ser Ser Ser Gly Pro Pro His Glu Thr Ala Ala Pro	
1255 1260 1265	
aaa cgc acc tat gac atg atg gag ggc cgt gta ggc agg act gtc acc	4014
Lys Arg Thr Tyr Asp Met Met Glu Gly Arg Val Gly Arg Thr Val Thr	
1270 1275 1280 1285	
tca gcc agc ata gag gga ctc atg ggc cgc gcc atc cct gag cag cac	4062
Ser Ala Ser Ile Glu Gly Leu Met Gly Arg Ala Ile Pro Glu Gln His	
1290 1295 1300	
agc ccc cac ctc aag gag cag cat cac atc cga ggc tcc atc acg caa	4110
Ser Pro His Leu Lys Glu Gln His His Ile Arg Gly Ser Ile Thr Gln	
1305 1310 1315	
ggc atc ccg agg tcc tat gtg gag gcg cag gag gac tac tta cgg cgg	4158
Gly Ile Pro Arg Ser Tyr Val Glu Ala Gln Glu Asp Tyr Leu Arg Arg	
1320 1325 1330	
gag gcc aag ctc ttg aag cga gaa ggg aca cca cct ccc cca cca cca	4206
Glu Ala Lys Leu Leu Lys Arg Glu Gly Thr Pro Pro Pro Pro Pro Pro	
1335 1340 1345	
cct cgg gac ctg act gag acc tac aag ccc cgg ccc ctg gac cct ctg	4254
Pro Arg Asp Leu Thr Glu Thr Tyr Lys Pro Arg Pro Leu Asp Pro Leu	
1350 1355 1360 1365	
ggt ccc ctg aag ctg aag ccg act cac gag ggt gtg gta gca act gtg	4302
Gly Pro Leu Lys Leu Lys Pro Thr His Glu Gly Val Val Ala Thr Val	
1370 1375 1380	
aag gag gcg ggc cgc tct atc cat gag atc ccg aga gag gag ctg cgc	4350
Lys Glu Ala Gly Arg Ser Ile His Glu Ile Pro Arg Glu Glu Leu Arg	
1385 1390 1395	
cgc aca cct gag cta ccc ctg gca cca cgg cct ctg aag gag ggt tcc	4398
Arg Thr Pro Glu Leu Pro Leu Ala Pro Arg Pro Leu Lys Glu Gly Ser	
1400 1405 1410	
atc acc cag ggc acc cca ctc aag tac gac tot ggg gca ccc tcc act	4446
Ile Thr Gln Gly Thr Pro Leu Lys Tyr Asp Ser Gly Ala Pro Ser Thr	
1415 1420 1425	
ggc acc aag aaa cac gac gtg cgc tcc atc atc ggc agc ccc ggc cgg	4494
Gly Thr Lys Lys His Asp Val Arg Ser Ile Ile Gly Ser Pro Gly Arg	
1430 1435 1440 1445	
cct ttc cct gcc ctg cac ccg ctg gac ata atg gct gac gcc cgg gca	4542
Pro Phe Pro Ala Leu His Pro Leu Asp Ile Met Ala Asp Ala Arg Ala	

1450										1455										1460										
ctg	gag	cgt	gcc	tgc	tat	gaa	gag	agt	ctg	aag	agc	cgg	tca	ggg	acc	4590														
Leu	Glu	Arg	Ala	Cys	Tyr	Glu	Glu	Ser	Leu	Lys	Ser	Arg	Ser	Gly	Thr															
1465					1470					1475																				
agc	agt	ggt	gca	ggg	ggc	tcc	atc	aca	cgt	ggg	gct	cca	gtc	gtc	gtg	4638														
Ser	Ser	Gly	Ala	Gly	Gly	Ser	Ile	Thr	Arg	Gly	Ala	Pro	Val	Val	Val															
1480					1485					1490																				
cct	gaa	ctg	ggc	aag	cca	cgg	caa	agc	cca	ctg	act	tac	gaa	gac	cac	4686														
Pro	Glu	Leu	Gly	Lys	Pro	Arg	Gln	Ser	Pro	Leu	Thr	Tyr	Glu	Asp	His															
1495					1500					1505																				
ggg	gca	ccc	ttc	acc	agt	cac	ctg	cca	cgt	ggc	tcc	cct	gtg	acc	acg	4734														
Gly	Ala	Pro	Phe	Thr	Ser	His	Leu	Pro	Arg	Gly	Ser	Pro	Val	Thr	Thr															
1510					1515					1520					1525															
agg	gag	ccc	acg	cca	cgc	ctt	cag	gaa	ggc	agc	ctc	cta	tcc	agc	aag	4782														
Arg	Glu	Pro	Thr	Pro	Arg	Leu	Gln	Glu	Gly	Ser	Leu	Leu	Ser	Ser	Lys															
1530					1535					1540																				
gcg	tcc	cag	gac	cgg	aag	ctg	aca	tct	aca	ccc	cgg	gag	atc	gcc	aag	4830														
Ala	Ser	Gln	Asp	Arg	Lys	Leu	Thr	Ser	Thr	Pro	Arg	Glu	Ile	Ala	Lys															
1545					1550					1555																				
tcc	cca	cac	agc	act	gtg	ccc	gag	cac	cac	cct	cac	ccc	atc	tcc	ccc	4878														
Ser	Pro	His	Ser	Thr	Val	Pro	Glu	His	His	Pro	His	Pro	Ile	Ser	Pro															
1560					1565					1570																				
tat	gag	cac	ttg	ctc	cgg	ggc	gtg	act	ggt	gtg	gac	ctg	tac	cgt	ggt	4926														
Tyr	Glu	His	Leu	Leu	Arg	Gly	Val	Thr	Gly	Val	Asp	Leu	Tyr	Arg	Gly															
1575					1580					1585																				
cac	atc	cca	ttg	gcc	ttt	gac	ccc	acc	tcc	ata	ccc	cga	ggg	atc	cct	4974														
His	Ile	Pro	Leu	Ala	Phe	Asp	Pro	Thr	Ser	Ile	Pro	Arg	Gly	Ile	Pro															
1590					1595					1600					1605															
ctg	gaa	gca	gca	gcc	gca	gcc	tac	tac	ctg	ccc	cgg	cac	ttg	gcc	ccc	5022														
Leu	Glu	Ala	Ala	Ala	Ala	Ala	Tyr	Tyr	Leu	Pro	Arg	His	Leu	Ala	Pro															
1610					1615					1620																				
agc	ccc	acc	tac	cca	cac	ctg	tac	cca	cct	tac	ctc	atc	cgc	ggc	tac	5070														
Ser	Pro	Thr	Tyr	Pro	His	Leu	Tyr	Pro	Pro	Tyr	Leu	Ile	Arg	Gly	Tyr															
1625					1630					1635																				
cct	gac	acg	gcg	gcc	ctg	gag	aac	cgc	cag	acc	atc	atc	aat	gac	tac	5118														
Pro	Asp	Thr	Ala	Ala	Leu	Glu	Asn	Arg	Gln	Thr	Ile	Ile	Asn	Asp	Tyr															
1640					1645					1650																				
atc	acc	tcg	cag	cag	atg	cac	cac	aac	gct	gcc	tcc	gcc	atg	gcc	cag	5166														
Ile	Thr	Ser	Gln	Gln	Met	His	His	Asn	Ala	Ala	Ser	Ala	Met	Ala	Gln															
1655					1660					1665																				
cgt	gct	gac	atg	ctg	agg	ggt	ctg	tca	ccg	cga	gag	tcc	tcg	ctg	gcc	5214														
Arg	Ala	Asp	Met	Leu	Arg	Gly	Leu	Ser	Pro	Arg	Glu	Ser	Ser	Leu	Ala															
1670					1675					1680					1685															
ctc	aat	tat	tcc	gct	ggc	cca	aga	ggc	att	atc	gac	ctg	tcc	caa	gtg	5262														
Leu	Asn	Tyr	Ser	Ala	Gly	Pro	Arg	Gly	Ile	Ile	Asp	Leu	Ser	Gln	Val															
1690					1695					170																				

cca cac ctg ccc gtg ctg gtg cca cca acg cca ggc acc cct gcc acc Pro His Leu Pro Val Leu Val Pro Pro Thr Pro Gly Thr Pro Ala Thr	5310
1705 1710 1715	
gcc atc gac cgc ctt gcc tac ctc ccc act gcg ccc cca ccc ttc agc Ala Ile Asp Arg Leu Ala Tyr Leu Pro Thr Ala Pro Pro Phe Ser	5358
1720 1725 1730	
agc cgc cac agt agc tca ccg ctg tcc cca gga ggc ccc act cac cta Ser Arg His Ser Ser Pro Leu Ser Pro Gly Gly Pro Thr His Leu	5406
1735 1740 1745	
gct aaa cca act gcc aca tct tca tcg gag cgg gaa cgg gaa cgt gag Ala Lys Pro Thr Ala Thr Ser Ser Ser Glu Arg Glu Arg Glu Arg Glu	5454
1750 1755 1760 1765	
cgg gaa cga gac aag tcc atc ctc acg tct acc act aca gtg gag cat Arg Glu Arg Asp Lys Ser Ile Leu Thr Ser Thr Thr Thr Val Glu His	5502
1770 1775 1780	
gca ccc atc tgg aga cct ggt acg gag cag agc agc ggg gct ggg ggc Ala Pro Ile Trp Arg Pro Gly Thr Glu Gln Ser Ser Gly Ala Gly Gly	5550
1785 1790 1795	
agc agc cgc ccc gcc tcc cac acc cac cag cac tcg ccc atc tcc ccc Ser Ser Arg Pro Ala Ser His Thr His Gln His Ser Pro Ile Ser Pro	5598
1800 1805 1810	
cgg acc cag gac gcc ttg cag cag agg ccc agt gtg ctg cac aac acg Arg Thr Gln Asp Ala Leu Gln Gln Arg Pro Ser Val Leu His Asn Thr	5646
1815 1820 1825	
agc atg aag ggc gtg gtc acc tcc gtg gaa ccc ggc acg ccc acg gtc Ser Met Lys Gly Val Val Thr Ser Val Glu Pro Gly Thr Pro Thr Val	5694
1830 1835 1840 1845	
ctg agg tcc acc tcc acc tct tcg cct gtc cgc cca gct gcc aca ttc Leu Arg Ser Thr Ser Thr Ser Ser Pro Val Arg Pro Ala Ala Thr Phe	5742
1850 1855 1860	
cca cct gcc acc cac tgc cca ctt ggt ggc acc ctt gaa ggg gtc tac Pro Pro Ala Thr His Cys Pro Leu Gly Gly Thr Leu Glu Gly Val Tyr	5790
1865 1870 1875	
cct acc ctc atg gag ccc gtc ctg tta ccc aag gag acc tct cgg gtc Pro Thr Leu Met Glu Pro Val Leu Leu Pro Lys Glu Thr Ser Arg Val	5838
1880 1885 1890	
gcc cgg ccc gag cgg ccc cgt gtg gac ggt ggc cat gcc ttc ctc acc Ala Arg Pro Glu Arg Pro Arg Val Asp Gly Gly His Ala Phe Leu Thr	5886
1895 1900 1905	
aaa ccc ccg gcc cgg gag ccc gcc tcc tca ccc agc aag agc tcc gag Lys Pro Pro Ala Arg Glu Pro Ala Ser Ser Pro Ser Lys Ser Ser Glu	5934
1910 1915 1920 1925	
ccc cga tcc cta gca ccc ccc agc tcc agc cac aca gcc atc gcc cgc Pro Arg Ser Leu Ala Pro Pro Ser Ser Ser His Thr Ala Ile Ala Arg	5982
1930 1935 1940	

acc cca gca aag agc ctt gca ccc cac cat gcc agt ccg gac ccg ccg	6030
Thr Pro Ala Lys Ser Leu Ala Pro His His Ala Ser Pro Asp Pro Pro	
1945 1950 1955	
ggg ccc acc tcg gcc tca gat ctg cac cga gaa aag act caa agt aaa	6078
Gly Pro Thr Ser Ala Ser Asp Leu His Arg Glu Lys Thr Gln Ser Lys	
1960 1965 1970	
ccc ttt tcc atc cag gaa ttg gaa ctc cgt tct ctg ggt tac cac agt	6126
Pro Phe Ser Ile Gln Glu Leu Glu Leu Arg Ser Leu Gly Tyr His Ser	
1975 1980 1985	
gga gct ggc tac agc ccc gat ggg gtg gag ccc atc agc ccg gtg agc	6174
Gly Ala Gly Tyr Ser Pro Asp Gly Val Glu Pro Ile Ser Pro Val Ser	
1990 1995 2000 2005	
tcc ccc agc ctg acc cac gac aag ggg ctc tcc aaa cct ctg gaa gag	6222
Ser Pro Ser Leu Thr His Asp Lys Gly Leu Ser Lys Pro Leu Glu Glu	
2010 2015 2020	
cta gag aag agc cac ttg gaa ggg gag ctg cgg cac aag cag cca ggc	6270
Leu Glu Lys Ser His Leu Glu Gly Glu Leu Arg His Lys Gln Pro Gly	
2025 2030 2035	
ccc atg aag ctc agc gcg gag gct gcc cat ctc cca cat ctg cgg cca	6318
Pro Met Lys Leu Ser Ala Glu Ala Ala His Leu Pro His Leu Arg Pro	
2040 2045 2050	
ctg ccc gag agc cag ccc tca tcc agc cca ctc ctc cag act gcc cca	6366
Leu Pro Glu Ser Gln Pro Ser Ser Ser Pro Leu Leu Gln Thr Ala Pro	
2055 2060 2065	
ggc atc aaa ggt cac cag agg gtg gtc acc ctg gct cag cac atc agc	6414
Gly Ile Lys Gly His Gln Arg Val Val Thr Leu Ala Gln His Ile Ser	
2070 2075 2080 2085	
gag gtc att acg cag gac tac acc cgg cac cac ccg cag cag ctc agt	6462
Glu Val Ile Thr Gln Asp Tyr Thr Arg His His Pro Gln Gln Leu Ser	
2090 2095 2100	
ggc ccc ctt ccc gcc cct ctc tac tcc ttt ccc gga gcc agc tgc cct	6510
Gly Pro Leu Pro Ala Pro Leu Tyr Ser Phe Pro Gly Ala Ser Cys Pro	
2105 2110 2115	
gtg ctg gat ctt cgc cgc cca ccc agt gac ctc tac ctc cca ccc ccc	6558
Val Leu Asp Leu Arg Arg Pro Pro Ser Asp Leu Tyr Leu Pro Pro Pro	
2120 2125 2130	
gac cat ggc acc cca gcc cgg gga tcc ccc cac agt gaa ggg ggc aaa	6606
Asp His Gly Thr Pro Ala Arg Gly Ser Pro His Ser Glu Gly Gly Lys	
2135 2140 2145	
agg tcc cca gaa ccc agc aaa aca tcg gtc ctg ggc agc agt gag gat	6654
Arg Ser Pro Glu Pro Ser Lys Thr Ser Val Leu Gly Ser Ser Glu Asp	
2150 2155 2160 2165	
gcc att gag cct gtg tcc cca cca gag ggc atg act gag cca gga cat	6702
Ala Ile Glu Pro Val Ser Pro Pro Glu Gly Met Thr Glu Pro Gly His	
2170 2175 2180	
gct cgg agc gct gtg tac cca ctg ctg tat cga gac ggg gaa cag ggc	6750

Ala Arg Ser Ala Val Tyr Pro Leu Leu Tyr Arg Asp Gly Glu Gln Gly	
2185 2190 2195	
gag ccc agg atg ggc tct aag tct cca ggc aac acc agc cag ccg cca	6798
Glu Pro Arg Met Gly Ser Lys Ser Pro Gly Asn Thr Ser Gln Pro Pro	
2200 2205 2210	
gcc ttc ttc agt aag ctg act gag agc aac tcc gcc atg gtg aag tcg	6846
Ala Phe Phe Ser Lys Leu Thr Glu Ser Asn Ser Ala Met Val Lys Ser	
2215 2220 2225	
aag aag cag gag atc aac aag aaa ctc aac acc cac aac cgg aac gag	6894
Lys Lys Gln Glu Ile Asn Lys Lys Leu Asn Thr His Asn Arg Asn Glu	
2230 2235 2240 2245	
cca gaa tac aat att ggc cag cct ggg acg gaa atc ttc aac atg ccc	6942
Pro Glu Tyr Asn Ile Gly Gln Pro Gly Thr Glu Ile Phe Asn Met Pro	
2250 2255 2260	
gcc atc act gga gca ggc ctt atg acc tgt aga agc cag gcg gtg caa	6990
Ala Ile Thr Gly Ala Gly Leu Met Thr Cys Arg Ser Gln Ala Val Gln	
2265 2270 2275	
gaa cac gcc agc acc aac atg ggg cta gag gcc att att aga aag gca	7038
Glu His Ala Ser Thr Asn Met Gly Leu Glu Ala Ile Ile Arg Lys Ala	
2280 2285 2290	
ctc atg ggt aaa tat gat cag tgg gaa gag ccc ccg ccg ctc ggc gcc	7086
Leu Met Gly Lys Tyr Asp Gln Trp Glu Glu Pro Pro Pro Leu Gly Ala	
2295 2300 2305	
aat gct ttt aac cct ctg aat gcc agc gcc agt ctg ccc gct gct gct	7134
Asn Ala Phe Asn Pro Leu Asn Ala Ser Ala Ser Leu Pro Ala Ala Ala	
2310 2315 2320 2325	
atg ccc ata acc act gct gac gga cgg agt gac cac gca ctc acc tcg	7182
Met Pro Ile Thr Thr Ala Asp Gly Arg Ser Asp His Ala Leu Thr Ser	
2330 2335 2340	
cca ggt gga ggt ggg aaa gcc aag gtc tct ggc aga cct agc agc cga	7230
Pro Gly Gly Gly Gly Lys Ala Lys Val Ser Gly Arg Pro Ser Ser Arg	
2345 2350 2355	
aaa gcc aag tcg cca gca cca ggc cta gcg tcc gga gac cga ccc cct	7278
Lys Ala Lys Ser Pro Ala Pro Gly Leu Ala Ser Gly Asp Arg Pro Pro	
2360 2365 2370	
tct gtc tcc tca gta cac tca gag ggg gac tgc aat cgc cga aca cca	7326
Ser Val Ser Ser Val His Ser Glu Gly Asp Cys Asn Arg Arg Thr Pro	
2375 2380 2385	
ctc acc aac cgt gtg tgg gag gac cgg ccc tca tct gca ggg tcc acg	7374
Leu Thr Asn Arg Val Trp Glu Asp Arg Pro Ser Ser Ala Gly Ser Thr	
2390 2395 2400 2405	
cca ttc ccc tac aac cct ttg att atg agg cta cag gca ggt gtc atg	7422
Pro Phe Pro Tyr Asn Pro Leu Ile Met Arg Leu Gln Ala Gly Val Met	
2410 2415 2420	
gcc tcc ccg ccc cca cct ggc ctt gcg gca ggc agc ggg ccc cta gct	7470
Ala Ser Pro Pro Pro Pro Gly Leu Ala Ala Gly Ser Gly Pro Leu Ala	

6798 6846 6894 6942 6990 7038 7086 7134 7182 7230 7278 7326 7374 7422 7470

2425 2430 2435

ggt ccc cac cac gcc tgg gat gag gag ccc aag cca ctg ctg tgt tca 7518
 Gly Pro His His Ala Trp Asp Glu Glu Pro Lys Pro Leu Leu Cys Ser
 2440 2445 2450

cag tat gag aca ctc tcg gac agc gag tgaccacgga ttggggggga 7565
 Gln Tyr Glu Thr Leu Ser Asp Ser Glu
 2455 2460

gcggtgccag gtcccgccaca aggcagaagc agcccagcat ggagcagaca gctgctgact 7625
 ccggagactg aggaaggagc ccctgagtct gcctgccgtc catccgtccg tccgtccact 7685
 catctgtcca tccagagctg gcatacctgcc tgtctaaagc cttaactaag actcccaccc 7745
 cgggctggcc ctgcgcagtg accttacact caggggatgt ttacctggtg ctcgagaggg 7805
 gagtggacag gaggggaggg acagcgggcc aggagggggg ggacagcaat cgtgtgtcag 7865
 tcgcactcgt gcatactgggg tcagcgggga cccaccacag gctgaccagg cacctccatg 7925
 ccaccgcctc gccccttacc gcatttgga ccaaagtcta actgaactct cgcgtgggtc 7985
 tgccccctctc tctgccccca gcccgcttgc ctctggacag acagacgttc ccagcttacc 8045
 ctgcccataat gctgtcatca tcgcagtctc caaaggccac ccagcccaca agactgggag 8105
 cccatcagac caggtgggtg acacaagggg cctggctggg gcacggatgc ttgcaggaac 8165
 tggaccgttt cccggcctgt tgctgtggca acgggagggg aggcacgtgt aaatggtgtt 8225
 ggcttacagg gtatatTTTT gataccttca atgagttaat tcagacgtct cacacaagga 8285
 aggactcgcc cgtgtttctc ccgtgtgtct ttggtctcta cctactgttt cagaggcacg 8345
 tgccagccaa ggcggtggcc caccatacgc aggacttggg ggtcaggggc tccggcacac 8405
 ggcaactgtgc ccttccccac accttacttc agcgaaatgg acttgatgcg tattctgtgg 8465
 ccgtctctctg tgcacggcgg cattctgtca ttacacatg ttgttccaat taaaaagcaa 8525
 atatactcaa gtgaaaaaa 8544

<210> 5
 <211> 2462
 <212> PRT
 <213> Mus musculus

<400> 5
 Met Ser Gly Ser Thr Gln Pro Val Ala Gln Thr Trp Arg Ala Ala Glu
 1 5 10 15
 Pro Arg Tyr Pro Pro His Gly Ile Ser Tyr Pro Val Gln Ile Ala Arg
 20 25 30
 Ser His Thr Asp Val Gly Leu Leu Glu Tyr Gln His His Pro Arg Asp
 35 40 45
 Tyr Thr Ser His Leu Ser Pro Gly Ser Ile Ile Gln Pro Gln Arg Arg

50					55					60					
Arg 65	Pro	Ser	Leu	Leu	Ser 70	Glu	Phe	Gln	Pro	Gly 75	Ser	Glu	Arg	Ser	Gln 80
Glu	Leu	His	Leu	Arg 85	Pro	Glu	Ser	Arg	Thr 90	Phe	Leu	Pro	Glu	Leu 95	Gly
Lys	Pro	Asp	Ile 100	Glu	Phe	Thr	Glu	Ser 105	Lys	Arg	Pro	Arg	Leu 110	Glu	Leu
Leu	Pro	Asp 115	Thr	Leu	Leu	Arg	Pro 120	Ser	Pro	Leu	Leu	Ala 125	Thr	Gly	Gln
Pro	Ser 130	Gly	Ser	Glu	Asp	Leu 135	Thr	Lys	Asp	Arg	Ser 140	Leu	Ala	Gly	Lys
Leu 145	Glu	Pro	Val	Ser 150	Pro	Pro	Ser	Pro	Pro	His 155	Ala	Asp	Pro	Glu	Leu 160
Glu	Leu	Ala	Pro	Ser 165	Arg	Leu	Ser	Lys	Glu 170	Glu	Leu	Ile	Gln	Asn 175	Arg
Leu	Asp	Arg	Val 180	Asp	Arg	Glu	Ile	Thr 185	Met	Val	Glu	Gln	Gln 190	Ile	Ser
Lys	Leu 195	Lys	Lys	Lys	Gln	Gln 200	Gln	Leu	Glu	Glu	Glu	Ala 205	Ala	Lys	Pro
Pro	Glu 210	Pro	Glu	Lys	Pro	Val 215	Ser	Pro	Pro	Pro	Ile 220	Glu	Ser	Lys	His
Arg 225	Ser	Leu	Val	Gln 230	Ile	Ile	Tyr	Asp	Glu	Asn 235	Arg	Lys	Lys	Ala	Glu 240
Ala	Ala	His	Arg	Ile 245	Leu	Glu	Gly	Leu	Gly 250	Pro	Gln	Val	Glu	Leu 255	Pro
Leu	Tyr	Asn	Gln 260	Pro	Ser	Asp	Thr	Arg 265	Gln	Tyr	His	Glu	Asn 270	Ile	Lys
Ile	Asn 275	Gln	Ala	Met	Arg	Lys	Lys 280	Leu	Ile	Leu	Tyr	Phe 285	Lys	Arg	Arg
Asn 290	His	Ala	Arg	Lys	Gln	Trp 295	Glu	Gln	Arg	Phe	Cys 300	Gln	Arg	Tyr	Asp
Gln 305	Leu	Met	Glu	Ala	Trp 310	Glu	Lys	Lys	Val	Glu 315	Arg	Ile	Glu	Asn	Asn 320
Pro	Arg	Arg	Arg	Ala 325	Lys	Glu	Ser	Lys	Val 330	Arg	Glu	Tyr	Tyr	Glu 335	Lys
Gln	Phe	Pro	Glu 340	Ile	Arg	Lys	Gln	Arg 345	Glu	Leu	Gln	Glu	Arg 350	Met	Gln
Ser	Arg	Val	Gly 355	Gln	Arg	Gly	Ser 360	Gly	Leu	Ser	Met	Ser 365	Ala	Ala	Arg
Ser 370	Glu	His	Glu	Val	Ser	Glu 375	Ile	Ile	Asp	Gly	Leu	Ser	Glu	Gln	Glu 380

Asn	Leu	Glu	Lys	Gln	Met	Arg	Gln	Leu	Ala	Val	Ile	Arg	His	Val	Val
385					390					395					400
Arg	Arg	Asp	Gln	Gln	Arg	Ile	Lys	Phe	Ile	Asn	Met	Asn	Gly	Leu	Met
				405					410					415	
Asp	Asp	Pro	Met	Lys	Val	Tyr	Lys	Asp	Arg	Gln	Val	Thr	Asn	Met	Trp
			420					425					430		
Ser	Glu	Gln	Glu	Arg	Asp	Thr	Phe	Arg	Glu	Lys	Phe	Met	Gln	His	Pro
			435				440					445			
Lys	Asn	Phe	Gly	Leu	Ile	Ala	Ser	Phe	Leu	Glu	Arg	Lys	Thr	Val	Ala
			450			455						460			
Glu	Cys	Val	Leu	Tyr	Tyr	Tyr	Leu	Thr	Lys	Lys	Asn	Glu	Asn	Tyr	Lys
465					470					475					480
Ser	Leu	Val	Arg	Arg	Ser	Tyr	Arg	Arg	Arg	Gly	Lys	Ser	Gln	Gln	Gln
				485					490					495	
Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Met	Ala	Arg	Ser	Ser
				500					505					510	
Gln	Glu	Glu	Lys	Glu	Glu	Lys	Glu	Lys	Glu	Lys	Glu	Ala	Asp	Lys	Glu
			515				520					525			
Glu	Glu	Lys	Gln	Asp	Ala	Glu	Asn	Glu	Lys	Glu	Glu	Leu	Ser	Lys	Glu
			530			535						540			
Lys	Thr	Asp	Asp	Thr	Ser	Gly	Glu	Asp	Asn	Asp	Glu	Lys	Glu	Ala	Val
545					550					555					560
Ala	Ser	Lys	Gly	Arg	Lys	Thr	Ala	Asn	Ser	Gln	Gly	Arg	Arg	Lys	Gly
				565					570					575	
Arg	Ile	Thr	Arg	Ser	Met	Ala	Asn	Glu	Ala	Asn	His	Glu	Glu	Thr	Ala
			580					585					590		
Thr	Pro	Gln	Gln	Ser	Ser	Glu	Leu	Ala	Ser	Met	Glu	Met	Asn	Glu	Ser
			595				600						605		
Ser	Arg	Trp	Thr	Glu	Glu	Glu	Met	Glu	Thr	Ala	Lys	Lys	Gly	Leu	Leu
						615					620				
Glu	His	Gly	Arg	Asn	Trp	Ser	Ala	Ile	Ala	Arg	Met	Val	Gly	Ser	Lys
625					630					635					640
Thr	Val	Ser	Gln	Cys	Lys	Asn	Phe	Tyr	Phe	Asn	Tyr	Lys	Lys	Arg	Gln
				645					650					655	
Asn	Leu	Asp	Glu	Ile	Leu	Gln	Gln	His	Lys	Leu	Lys	Met	Glu	Lys	Glu
			660					665					670		
Arg	Asn	Ala	Arg	Arg	Lys	Lys	Lys	Lys	Thr	Pro	Ala	Ala	Ala	Ser	Glu
			675				680					685			
Glu	Thr	Ala	Phe	Pro	Pro	Ala	Ala	Glu	Asp	Glu	Glu	Met	Glu	Ala	Ser
						695						700			

Gly 705	Ala	Ser	Ala	Asn 710	Glu	Glu	Glu	Leu	Ala	Glu 715	Glu	Ala	Glu	Ala	Ser 720
Gln	Ala	Ser	Gly 725	Asn	Glu	Val	Pro	Arg	Val 730	Gly	Glu	Cys	Ser	Gly 735	Pro
Ala	Ala	Val	Asn 740	Asn	Ser	Ser	Asp	Thr 745	Glu	Ser	Val	Pro	Ser 750	Pro	Arg
Ser	Glu	Ala 755	Thr	Lys	Asp	Thr	Gly 760	Pro	Lys	Pro	Thr	Gly 765	Thr	Glu	Ala
Leu	Pro 770	Ala	Ala	Thr	Gln	Pro 775	Pro	Val	Pro	Pro	Pro 780	Glu	Glu	Pro	Ala
Ala 785	Ala	Pro	Ala	Glu	Pro 790	Ser	Pro	Val	Pro	Asp 795	Ala	Ser	Gly	Pro	Pro 800
Ser	Pro	Glu	Pro	Ser 805	Pro	Ser	Pro	Ala	Ala 810	Pro	Pro	Ala	Thr	Val 815	Asp
Lys	Asp	Glu	Gln 820	Glu	Ala	Pro	Ala	Ala 825	Pro	Ala	Pro	Gln	Thr 830	Glu	Asp
Ala	Lys	Glu 835	Gln	Lys	Ser	Glu	Ala 840	Glu	Glu	Ile	Asp	Val 845	Gly	Lys	Pro
Glu	Glu 850	Pro	Glu	Ala	Ser	Glu 855	Glu	Pro	Pro	Glu	Ser	Val 860	Lys	Ser	Asp
His 865	Lys	Glu	Glu	Thr	Glu 870	Glu	Glu	Pro	Glu	Asp 875	Lys	Ala	Lys	Gly	Thr 880
Glu	Ala	Ile	Glu	Thr 885	Val	Ser	Glu	Ala	Pro 890	Leu	Lys	Val	Glu	Glu 895	Ala
Gly	Ser	Lys	Ala 900	Ala	Val	Thr	Lys	Gly 905	Ser	Ser	Ser	Gly	Ala 910	Thr	Gln
Asp	Ser	Asp 915	Ser	Ser	Ala	Thr	Cys 920	Ser	Ala	Asp	Glu	Val 925	Asp	Glu	Pro
Glu	Gly 930	Gly	Asp	Lys	Gly	Arg 935	Leu	Leu	Ser	Pro	Arg 940	Pro	Ser	Leu	Leu
Thr 945	Pro	Ala	Gly	Asp	Pro 950	Arg	Ala	Ser	Thr	Ser 955	Pro	Gln	Lys	Pro	Leu 960
Asp	Leu	Lys	Gln	Leu 965	Lys	Gln	Arg	Ala	Ala 970	Ala	Ile	Pro	Pro	Ile 975	Val
Thr	Lys	Val	His 980	Glu	Pro	Pro	Arg	Glu 985	Asp	Thr	Val	Pro	Pro 990	Lys	Pro
Val	Pro	Pro 995	Val	Pro	Pro	Pro	Thr 1000	Gln	His	Leu	Gln	Pro 1005	Glu	Gly	Asp
Val 1010	Ser	Gln	Gln	Ser	Gly	Gly 1015	Ser	Pro	Arg	Gly	Lys 1020	Ser	Arg	Ser	Pro
Val	Pro	Pro	Ala	Glu	Lys	Glu	Ala	Glu	Lys	Pro	Ala	Phe	Phe	Pro	Ala

025	1030							1035							1040			
Phe	Pro	Thr	Glu	Gly	Gln	Ser	Tyr	Arg	Leu	Ser	Pro	His	Ala	Gly	His			
				1045					1050					1055				
Arg	Leu	Pro	Ser	His	Pro	Pro	Arg	Glu	Val	Ile	Lys	Thr	Ser	Thr	Arg			
			1060					1065					1070					
Ala	Asp	Pro	Leu	Phe	Ser	Tyr	Thr	Pro	Pro	Gly	His	Pro	Leu	Pro	Leu			
		1075					1080					1085						
Gly	Leu	His	Asp	Ser	Ala	Arg	Pro	Val	Leu	Pro	Arg	Pro	Pro	Ile	Ser			
	1090					1095					1100							
Asn	Pro	Pro	Pro	Leu	Ile	Ser	Ser	Ala	Lys	His	Pro	Gly	Val	Leu	Glu			
105					1110					1115					1120			
Arg	Gln	Leu	Gly	Ala	Ile	Ser	Gln	Gly	Met	Ser	Val	Gln	Leu	Arg	Val			
				1125					1130					1135				
Pro	His	Ser	Glu	His	Ala	Lys	Pro	Met	Gly	Pro	Leu	Thr	Met	Glu	Leu			
			1140					1145					1150					
Pro	Leu	Ala	Val	Asp	Pro	Lys	Lys	Leu	Gly	Thr	Ala	Leu	Ala	Pro	Pro			
		1155					1160					1165						
Pro	Val	Glu	Ala	Ser	Pro	Arg	Ala	Ser	Gln	Tyr	Pro	Gly	Cys	Arg	Arg			
	1170					1175					1180							
Pro	Gln	Leu	Gln	Arg	Leu	Tyr	His	Pro	Arg	Thr	Pro	Ala	Asp	Val	Leu			
185					1190					1195					1200			
Tyr	Lys	Gly	Thr	Ile	Ser	Arg	Ile	Val	Gly	Glu	Asp	Ser	Pro	Ser	Arg			
				1205					1210					1215				
Leu	Asp	Arg	Ala	Arg	Glu	Asp	Thr	Leu	Pro	Lys	Gly	His	Val	Ile	Tyr			
			1220					1225					1230					
Glu	Gly	Lys	Lys	Gly	His	Val	Leu	Ser	Tyr	Glu	Gly	Gly	Met	Ser	Val			
		1235					1240					1245						
Ser	Gln	Cys	Ser	Lys	Glu	Asp	Gly	Arg	Ser	Ser	Ser	Gly	Pro	Pro	His			
	1250					1255					1260							
Glu	Thr	Ala	Ala	Pro	Lys	Arg	Thr	Tyr	Asp	Met	Met	Glu	Gly	Arg	Val			
265					1270					1275					1280			
Gly	Arg	Thr	Val	Thr	Ser	Ala	Ser	Ile	Glu	Gly	Leu	Met	Gly	Arg	Ala			
				1285					1290					1295				
Ile	Pro	Glu	Gln	His	Ser	Pro	His	Leu	Lys	Glu	Gln	His	His	Ile	Arg			
			1300					1305					1310					
Gly	Ser	Ile	Thr	Gln	Gly	Ile	Pro	Arg	Ser	Tyr	Val	Glu	Ala	Gln	Glu			
		1315					1320					1325						
Asp	Tyr	Leu	Arg	Arg	Glu	Ala	Lys	Leu	Leu	Lys	Arg	Glu	Gly	Thr	Pro			
	1330					1335					1340							
Pro	Pro	Pro	Pro	Pro	Pro	Arg	Asp	Leu	Thr	Glu	Thr	Tyr	Lys	Pro	Arg			
345					1350					1355					1360			

Pro Leu Asp Pro Leu Gly Pro Leu Lys Leu Lys Pro Thr His Glu Gly
1365 1370 1375

Val Val Ala Thr Val Lys Glu Ala Gly Arg Ser Ile His Glu Ile Pro
1380 1385 1390

Arg Glu Glu Leu Arg Arg Thr Pro Glu Leu Pro Leu Ala Pro Arg Pro
1395 1400 1405

Leu Lys Glu Gly Ser Ile Thr Gln Gly Thr Pro Leu Lys Tyr Asp Ser
1410 1415 1420

Gly Ala Pro Ser Thr Gly Thr Lys Lys His Asp Val Arg Ser Ile Ile
425 1430 1435 1440

Gly Ser Pro Gly Arg Pro Phe Pro Ala Leu His Pro Leu Asp Ile Met
1445 1450 1455

Ala Asp Ala Arg Ala Leu Glu Arg Ala Cys Tyr Glu Glu Ser Leu Lys
1460 1465 1470

Ser Arg Ser Gly Thr Ser Ser Gly Ala Gly Gly Ser Ile Thr Arg Gly
1475 1480 1485

Ala Pro Val Val Val Pro Glu Leu Gly Lys Pro Arg Gln Ser Pro Leu
1490 1495 1500

Thr Tyr Glu Asp His Gly Ala Pro Phe Thr Ser His Leu Pro Arg Gly
505 1510 1515 1520

Ser Pro Val Thr Thr Arg Glu Pro Thr Pro Arg Leu Gln Glu Gly Ser
1525 1530 1535

Leu Leu Ser Ser Lys Ala Ser Gln Asp Arg Lys Leu Thr Ser Thr Pro
1540 1545 1550

Arg Glu Ile Ala Lys Ser Pro His Ser Thr Val Pro Glu His His Pro
1555 1560 1565

His Pro Ile Ser Pro Tyr Glu His Leu Leu Arg Gly Val Thr Gly Val
1570 1575 1580

Asp Leu Tyr Arg Gly His Ile Pro Leu Ala Phe Asp Pro Thr Ser Ile
585 1590 1595 1600

Pro Arg Gly Ile Pro Leu Glu Ala Ala Ala Ala Tyr Tyr Leu Pro
1605 1610 1615

Arg His Leu Ala Pro Ser Pro Thr Tyr Pro His Leu Tyr Pro Pro Tyr
1620 1625 1630

Leu Ile Arg Gly Tyr Pro Asp Thr Ala Ala Leu Glu Asn Arg Gln Thr
1635 1640 1645

Ile Ile Asn Asp Tyr Ile Thr Ser Gln Gln Met His His Asn Ala Ala
1650 1655 1660

Ser Ala Met Ala Gln Arg Ala Asp Met Leu Arg Gly Leu Ser Pro Arg
665 1670 1675 1680

1660
1655
1650
1645
1640
1635
1630
1625
1620
1615
1610
1605
1600
1595
1590
1585
1580
1575
1570
1565
1560
1555
1550
1545
1540
1535
1530
1525
1520
1515
1510
1505
1500
1495
1490
1485
1480
1475
1470
1465
1460
1455
1450
1445
1440
1435
1430
1425
1420
1415
1410
1405
1400
1395
1390
1385
1380
1375
1370
1365

Glu Ser Ser Leu Ala Leu Asn Tyr Ser Ala Gly Pro Arg Gly Ile Ile
1685 1690 1695

Asp Leu Ser Gln Val Pro His Leu Pro Val Leu Val Pro Pro Thr Pro
1700 1705 1710

Gly Thr Pro Ala Thr Ala Ile Asp Arg Leu Ala Tyr Leu Pro Thr Ala
1715 1720 1725

Pro Pro Pro Phe Ser Ser Arg His Ser Ser Ser Pro Leu Ser Pro Gly
1730 1735 1740

Gly Pro Thr His Leu Ala Lys Pro Thr Ala Thr Ser Ser Ser Glu Arg
745 1750 1755 1760

Glu Arg Glu Arg Glu Arg Glu Arg Asp Lys Ser Ile Leu Thr Ser Thr
1765 1770 1775

Thr Thr Val Glu His Ala Pro Ile Trp Arg Pro Gly Thr Glu Gln Ser
1780 1785 1790

Ser Gly Ala Gly Gly Ser Ser Arg Pro Ala Ser His Thr His Gln His
1795 1800 1805

Ser Pro Ile Ser Pro Arg Thr Gln Asp Ala Leu Gln Gln Arg Pro Ser
1810 1815 1820

Val Leu His Asn Thr Ser Met Lys Gly Val Val Thr Ser Val Glu Pro
825 1830 1835 1840

Gly Thr Pro Thr Val Leu Arg Ser Thr Ser Thr Ser Ser Pro Val Arg
1845 1850 1855

Pro Ala Ala Thr Phe Pro Pro Ala Thr His Cys Pro Leu Gly Gly Thr
1860 1865 1870

Leu Glu Gly Val Tyr Pro Thr Leu Met Glu Pro Val Leu Leu Pro Lys
1875 1880 1885

Glu Thr Ser Arg Val Ala Arg Pro Glu Arg Pro Arg Val Asp Gly Gly
1890 1895 1900

His Ala Phe Leu Thr Lys Pro Pro Ala Arg Glu Pro Ala Ser Ser Pro
905 1910 1915 1920

Ser Lys Ser Ser Glu Pro Arg Ser Leu Ala Pro Pro Ser Ser Ser His
1925 1930 1935

Thr Ala Ile Ala Arg Thr Pro Ala Lys Ser Leu Ala Pro His His Ala
1940 1945 1950

Ser Pro Asp Pro Pro Gly Pro Thr Ser Ala Ser Asp Leu His Arg Glu
1955 1960 1965

Lys Thr Gln Ser Lys Pro Phe Ser Ile Gln Glu Leu Glu Leu Arg Ser
1970 1975 1980

Leu Gly Tyr His Ser Gly Ala Gly Tyr Ser Pro Asp Gly Val Glu Pro
985 1990 1995 2000

Ile Ser Pro Val Ser Ser Pro Ser Leu Thr His Asp Lys Gly Leu Ser

1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000

2005					2010					2015							
Lys	Pro	Leu	Glu	Glu	Leu	Glu	Lys	Ser	His	Leu	Glu	Gly	Glu	Leu	Arg		
2020					2025					2030							
His	Lys	Gln	Pro	Gly	Pro	Met	Lys	Leu	Ser	Ala	Glu	Ala	Ala	His	Leu		
2035					2040					2045							
Pro	His	Leu	Arg	Pro	Leu	Pro	Glu	Ser	Gln	Pro	Ser	Ser	Ser	Pro	Leu		
2050					2055					2060							
Leu	Gln	Thr	Ala	Pro	Gly	Ile	Lys	Gly	His	Gln	Arg	Val	Val	Thr	Leu		
065					2070				2075				2080				
Ala	Gln	His	Ile	Ser	Glu	Val	Ile	Thr	Gln	Asp	Tyr	Thr	Arg	His	His		
2085					2090					2095							
Pro	Gln	Gln	Leu	Ser	Gly	Pro	Leu	Pro	Ala	Pro	Leu	Tyr	Ser	Phe	Pro		
2100					2105					2110							
Gly	Ala	Ser	Cys	Pro	Val	Leu	Asp	Leu	Arg	Arg	Pro	Pro	Ser	Asp	Leu		
2115					2120					2125							
Tyr	Leu	Pro	Pro	Pro	Asp	His	Gly	Thr	Pro	Ala	Arg	Gly	Ser	Pro	His		
2130					2135					2140							
Ser	Glu	Gly	Gly	Lys	Arg	Ser	Pro	Glu	Pro	Ser	Lys	Thr	Ser	Val	Leu		
145					2150				2155				2160				
Gly	Ser	Ser	Glu	Asp	Ala	Ile	Glu	Pro	Val	Ser	Pro	Pro	Glu	Gly	Met		
2165					2170					2175							
Thr	Glu	Pro	Gly	His	Ala	Arg	Ser	Ala	Val	Tyr	Pro	Leu	Leu	Tyr	Arg		
2180					2185					2190							
Asp	Gly	Glu	Gln	Gly	Glu	Pro	Arg	Met	Gly	Ser	Lys	Ser	Pro	Gly	Asn		
2195					2200					2205							
Thr	Ser	Gln	Pro	Pro	Ala	Phe	Phe	Ser	Lys	Leu	Thr	Glu	Ser	Asn	Ser		
2210					2215					2220							
Ala	Met	Val	Lys	Ser	Lys	Lys	Gln	Glu	Ile	Asn	Lys	Lys	Leu	Asn	Thr		
2225					2230					2235					2240		
His	Asn	Arg	Asn	Glu	Pro	Glu	Tyr	Asn	Ile	Gly	Gln	Pro	Gly	Thr	Glu		
2245					2250					2255							
Ile	Phe	Asn	Met	Pro	Ala	Ile	Thr	Gly	Ala	Gly	Leu	Met	Thr	Cys	Arg		
2260					2265					2270							
Ser	Gln	Ala	Val	Gln	Glu	His	Ala	Ser	Thr	Asn	Met	Gly	Leu	Glu	Ala		
2275					2280					2285							
Ile	Ile	Arg	Lys	Ala	Leu	Met	Gly	Lys	Tyr	Asp	Gln	Trp	Glu	Glu	Pro		
2290					2295					2300							
Pro	Pro	Leu	Gly	Ala	Asn	Ala	Phe	Asn	Pro	Leu	Asn	Ala	Ser	Ala	Ser		
305					2310					2315					2320		
Leu	Pro	Ala	Ala	Ala	Met	Pro	Ile	Thr	Thr	Ala	Asp	Gly	Arg	Ser	Asp		
2325					2330					2335							

2005 2010 2015

<400>	6																
atg	tca	gga	tcc	aca	cag	cct	gtg	gca	cag	aca	tgg	cgg	gct	gct	gag		48
Met	Ser	Gly	Ser	Thr	Gln	Pro	Val	Ala	Gln	Thr	Trp	Arg	Ala	Ala	Glu		
1				5					10					15			
ccc	cgc	tac	cca	ccc	cat	ggc	atc	tcc	tac	ccg	gtg	cag	ata	gcc	cgg		96
Pro	Arg	Tyr	Pro	Pro	His	Gly	Ile	Ser	Tyr	Pro	Val	Gln	Ile	Ala	Arg		
			20					25					30				
tcc	cac	acg	gac	gtg	ggg	ctg	ctt	gag	tac	caa	cac	cac	ccc	cgt	gac		144
Ser	His	Thr	Asp	Val	Gly	Leu	Leu	Glu	Tyr	Gln	His	His	Pro	Arg	Asp		
		35				40						45					
tac	acc	tca	cac	ctg	tca	ccc	ggt	tcc	atc	atc	cag	cca	cag	agg	agg		192
Tyr	Thr	Ser	His	Leu	Ser	Pro	Gly	Ser	Ile	Ile	Gln	Pro	Gln	Arg	Arg		
	50					55					60						
cgg	ccc	tca	ctg	ctg	tca	gag	ttc	cag	cct	ggg	agt	gaa	cgg	tct	cag		240
Arg	Pro	Ser	Leu	Leu	Ser	Glu	Phe	Gln	Pro	Gly	Ser	Glu	Arg	Ser	Gln		
65					70					75					80		
gag	ctc	cac	ctg	cgc	cct	gag	tcc	cgc	acg	ttc	ctg	cct	gag	ctg	ggc		288
Glu	Leu	His	Leu	Arg	Pro	Glu	Ser	Arg	Thr	Phe	Leu	Pro	Glu	Leu	Gly		
				85				90						95			
aaq	ccc	gac	ata	gaa	ttc	acc	gag	agc	aag	cgc	ccc	cgc	ctg	gag	cta		336

Lys	Pro	Asp	Ile	Glu	Phe	Thr	Glu	Ser	Lys	Arg	Pro	Arg	Leu	Glu	Leu	
			100				105						110			
cta	ccc	gat	acc	ctg	ctg	cgc	cca	tca	ccc	ctg	ctg	gcc	act	ggg	cag	384
Leu	Pro	Asp	Thr	Leu	Leu	Arg	Pro	Ser	Pro	Leu	Leu	Ala	Thr	Gly	Gln	
			115				120						125			
ccg	agt	ggg	tct	gaa	gac	ctt	acc	aag	gac	cgt	agc	ctg	gca	ggc	aag	432
Pro	Ser	Gly	Ser	Glu	Asp	Leu	Thr	Lys	Asp	Arg	Ser	Leu	Ala	Gly	Lys	
			130				135						140			
ctg	gag	cct	gtg	tca	cct	ccc	agt	ccc	ccg	cac	gct	gac	cct	gag	cta	480
Leu	Glu	Pro	Val	Ser	Pro	Pro	Ser	Pro	Pro	His	Ala	Asp	Pro	Glu	Leu	
			145				150						155	160		
gag	ctg	gcg	cca	tct	cga	ctg	tcc	aag	gag	gag	ctg	atc	cag	aac	aga	528
Glu	Leu	Ala	Pro	Ser	Arg	Leu	Ser	Lys	Glu	Glu	Leu	Ile	Gln	Asn	Arg	
			165				170						175			
ttg	gac	cgc	gtg	gac	cgt	gag	atc	acc	atg	gta	gag	cag	cag	atc	tcc	576
Leu	Asp	Arg	Val	Asp	Arg	Glu	Ile	Thr	Met	Val	Glu	Gln	Gln	Ile	Ser	
			180				185						190			
aag	ctg	aag	aag	aag	cag	caa	cag	ttg	gag	gag	gag	gcc	gcc	aag	ccg	624
Lys	Leu	Lys	Lys	Lys	Gln	Gln	Gln	Leu	Glu	Glu	Glu	Ala	Ala	Lys	Pro	
			195				200						205			
ccc	gaa	ccc	gag	aag	cct	gtg	tcg	cca	cca	ccc	ata	gaa	tca	aag	cac	672
Pro	Glu	Pro	Glu	Lys	Pro	Val	Ser	Pro	Pro	Pro	Ile	Glu	Ser	Lys	His	
			210				215						220			
cga	agc	ctg	gtc	cag	atc	atc	tac	gat	gag	aac	cgg	aag	aaa	gcc	gaa	720
Arg	Ser	Leu	Val	Gln	Ile	Ile	Tyr	Asp	Glu	Asn	Arg	Lys	Lys	Ala	Glu	
			225				230						235	240		
gcc	gca	cac	cgg	atc	cta	gaa	ggc	ctg	ggg	ccc	cag	gtg	gag	ctg	cct	768
Ala	Ala	His	Arg	Ile	Leu	Glu	Gly	Leu	Gly	Pro	Gln	Val	Glu	Leu	Pro	
			245				250						255			
ctg	tac	aac	cag	ccg	tct	gac	aca	cgc	cag	tac	cat	gaa	aac	atc	aaa	816
Leu	Tyr	Asn	Gln	Pro	Ser	Asp	Thr	Arg	Gln	Tyr	His	Glu	Asn	Ile	Lys	
			260				265						270			
ata	aac	cag	gcg	atg	cgg	aag	aag	ctg	atc	ttg	tac	ttt	aag	cgg	agg	864
Ile	Asn	Gln	Ala	Met	Arg	Lys	Lys	Leu	Ile	Leu	Tyr	Phe	Lys	Arg	Arg	
			275				280						285			
aac	cac	gcg	cgc	aag	cag	tgg	gaa	cag	cgc	ttc	tgc	cag	cgc	tat	gac	912
Asn	His	Ala	Arg	Lys	Gln	Trp	Glu	Gln	Arg	Phe	Cys	Gln	Arg	Tyr	Asp	
			290				295						300			
cag	ctc	atg	gag	gcg	tgg	gag	aag	aag	gta	gag	cgc	ata	gag	aac	aat	960
Gln	Leu	Met	Glu	Ala	Trp	Glu	Lys	Lys	Val	Glu	Arg	Ile	Glu	Asn	Asn	
			305				310						315	320		
ccg	cga	agg	agg	gcc	aag	gag	agc	aag	gtg	agg	gag	tac	tac	gag	aaa	1008
Pro	Arg	Arg	Arg	Ala	Lys	Glu	Ser	Lys	Val	Arg	Glu	Tyr	Tyr	Glu	Lys	
			325				3									

340						345						350						
agc Ser	agg Arg	gtg Val 355	ggc Gly	cag Gln	cgt Arg	ggc Gly	agt Ser 360	ggg Gly	ctc Leu	tcc Ser	atg Met	tcg Ser 365	gct Ala	gcc Ala	cgc Arg	1104		
agt Ser	gag Glu 370	cat His	gag Glu	gtt Val	tct Ser	gag Glu 375	atc Ile	att Ile	gat Asp	ggc Gly	ttg Leu 380	tct Ser	gag Glu	cag Gln	gag Glu	1152		
aac Asn 385	ctg Leu	gag Glu	aag Lys	cag Gln	atg Met 390	cgc Arg	cag Gln	ctg Leu	gcc Ala	gtg Val 395	atc Ile	cgc Arg	cat His	gtt Val	gta Val 400	1200		
cga Arg	cgc Arg	gac Asp	cag Gln	cag Gln 405	agg Arg	atc Ile	aag Lys	ttc Phe	atc Ile 410	aac Asn	atg Met	aat Asn	gga Gly	ctc Leu 415	atg Met	1248		
gat Asp	gac Asp	ccc Pro 420	atg Met	aag Lys	gtc Val	tac Tyr	aag Lys	gac Asp 425	cgt Arg	cag Gln	gtt Val	acc Thr 430	aac Asn	atg Met	tgg Trp	1296		
agc Ser	gag Glu 435	cag Gln	gag Glu	agg Arg	gac Asp	acc Thr	ttc Phe 440	cgt Arg	gag Glu	aag Lys	ttt Phe 445	atg Met 445	cag Gln	cac His	cct Pro	1344		
aag Lys 450	aac Asn	ttt Phe	ggc Gly	ctg Leu	att Ile	gcc Ala 455	tca Ser	ttc Phe	ctg Leu	gag Glu 460	aga Arg	aag Lys	acg Thr	gtc Val	gct Ala	1392		
gag Glu 465	tgt Cys	gtc Val	ctc Leu	tat Tyr	tac Tyr 470	tac Tyr	ctg Leu	acc Thr	aag Lys	aag Lys 475	aat Asn	gaa Glu	aat Asn	tac Tyr	aag Lys 480	1440		
agc Ser	ttg Leu	gtg Val	agg Arg	cgg Arg 485	agc Ser	tat Tyr	cgg Arg	cgc Arg	cgt Arg 490	ggc Gly	aag Lys	agc Ser	cag Gln	cag Gln 495	cag Gln	1488		
cag Gln	cag Gln	cag Gln 500	caa Gln	caa Gln	cag Gln	cag Gln	cag Gln 505	cag Gln	cag Gln	cag Gln	atg Met	gca Ala	cgg Arg 510	agc Ser	agc Ser	1536		
cag Gln	gag Glu 515	gag Glu	aag Lys	gag Glu	gag Glu	aag Lys	gag Glu 520	aag Lys	gag Glu	aag Lys	gag Glu 525	gcc Ala	gac Asp	aag Lys	gag Glu	1584		
gaa Glu 530	gag Glu	aag Lys	cag Gln	gat Asp	gcg Ala	gag Glu 535	aac Asn	gag Glu	aag Lys	gaa Glu	gaa Glu 540	ctc Leu	agc Ser	aag Lys	gag Glu	1632		
aag Lys 545	aca Thr	gac Asp	gac Asp	act Thr	tct Ser 550	ggc Gly	gag Glu	gac Asp	aac Asn	gat Asp 555	gag Glu	aaa Lys	gag Glu	gcc Ala	gtg Val 560	1680		
gcc Ala	tcc Ser	aaa Lys	ggc Gly	cgc Arg 565	aaa Lys	act Thr	gcc Ala	aac Asn	agc Ser 570	caa Gln	ggc Gly	cgc Arg	cgc Arg	aaa Lys 575	ggc Gly	1728		
cgt Arg	atc Ile	acg Thr 580	cgc Arg	tcc Ser	atg Met	gcc Ala	aac Asn	gag Glu 585	gcc Ala	aac Asn	cat His	gag Glu 590	gag Glu	aca Thr	gcc Ala	1776		

acc cca cag caa agt tca gag ctg gct tcc atg gag atg aac gag agt	1824
Thr Pro Gln Gln Ser Ser Glu Leu Ala Ser Met Glu Met Asn Glu Ser	
595 600 605	
tct cgc tgg act gag gaa gag atg gag aca gca aag aaa ggc ctc ctg	1872
Ser Arg Trp Thr Glu Glu Glu Met Glu Thr Ala Lys Lys Gly Leu Leu	
610 615 620	
gaa cat ggg agg aac tgg tca gcc att gcc cgc atg gtg ggc tcc aag	1920
Glu His Gly Arg Asn Trp Ser Ala Ile Ala Arg Met Val Gly Ser Lys	
625 630 635 640	
acc gtg tcc cag tgt aag aac ttc tac ttc aac tac aag aag agg cag	1968
Thr Val Ser Gln Cys Lys Asn Phe Tyr Phe Asn Tyr Lys Lys Arg Gln	
645 650 655	
aac ctg gac gaa atc ctt cag cag cac aag cta aag atg gag aag gag	2016
Asn Leu Asp Glu Ile Leu Gln Gln His Lys Leu Lys Met Glu Lys Glu	
660 665 670	
agg aac gct cgg agg aag aag aag aag acc cca gct gcg gcg agc gag	2064
Arg Asn Ala Arg Arg Lys Lys Lys Lys Thr Pro Ala Ala Ala Ser Glu	
675 680 685	
gag aca gcc ttc cca cct gcc gct gag gac gaa gag atg gaa gca tca	2112
Glu Thr Ala Phe Pro Pro Ala Ala Glu Asp Glu Glu Met Glu Ala Ser	
690 695 700	
ggc gca agt gcc aat gag gaa gag ctg gcg gag gag gca gaa gcc tca	2160
Gly Ala Ser Ala Asn Glu Glu Glu Leu Ala Glu Glu Ala Glu Ala Ser	
705 710 715 720	
cag gcc tct ggg aat gag gtt ccc aga gtt ggg gag tgc agt ggc cca	2208
Gln Ala Ser Gly Asn Glu Val Pro Arg Val Gly Glu Cys Ser Gly Pro	
725 730 735	
gct gct gtc aac aac agc tct gat act gag agt gtc cca tcc ccg cgt	2256
Ala Ala Val Asn Asn Ser Ser Asp Thr Glu Ser Val Pro Ser Pro Arg	
740 745 750	
tca gaa gcc acg aag gac act ggg cct aaa ccc act ggc act gaa gca	2304
Ser Glu Ala Thr Lys Asp Thr Gly Pro Lys Pro Thr Gly Thr Glu Ala	
755 760 765	
ttg ccc gct gcc acc cag cca cct gtt cct cct cca gaa gaa ccg gca	2352
Leu Pro Ala Ala Thr Gln Pro Pro Val Pro Pro Pro Glu Glu Pro Ala	
770 775 780	
gca gcc cct gct gag ccc tcc cca gtc cct gat gcc agt ggc cca cca	2400
Ala Ala Pro Ala Glu Pro Ser Pro Val Pro Asp Ala Ser Gly Pro Pro	
785 790 795 800	
tcc cca gag cct tcc cca tca cct gcc gca ccc ccg gct act gtg gac	2448
Ser Pro Glu Pro Ser Pro Ser Pro Ala Ala Pro Pro Ala Thr Val Asp	
805 810 815	
aag gat gaa caa gaa gcc ccg gct gct cca gct ccc cag aca gaa gat	2496
Lys Asp Glu Gln Glu Ala Pro Ala Ala Pro Ala Pro Gln Thr Glu Asp	
820 825 830	

102220-102230

gcc aag gag cag aag tct gag gcc gag gag atc gat gtg gga aag cca	2544
Ala Lys Glu Gln Lys Ser Glu Ala Glu Glu Ile Asp Val Gly Lys Pro	
835 840 845	
gag gag ccc gag gcc tct gag gag ccc ccg gag agt gta aag agt gac	2592
Glu Glu Pro Glu Ala Ser Glu Glu Pro Pro Glu Ser Val Lys Ser Asp	
850 855 860	
cac aag gag gag acc gag gaa gag cct gaa gac aaa gcc aag ggc aca	2640
His Lys Glu Glu Thr Glu Glu Glu Pro Glu Asp Lys Ala Lys Gly Thr	
865 870 875 880	
gag gcc att gaa act gtg tct gag gca cca ctt aag gtg gag gag gct	2688
Glu Ala Ile Glu Thr Val Ser Glu Ala Pro Leu Lys Val Glu Glu Ala	
885 890 895	
ggt agc aag gca gct gtg acc aag ggt tcc agc tca ggt gcc acc cag	2736
Gly Ser Lys Ala Ala Val Thr Lys Gly Ser Ser Ser Gly Ala Thr Gln	
900 905 910	
gac agt gac tcc agt gcc acc tgc agt gcc gat gag gtg gac gaa ccc	2784
Asp Ser Asp Ser Ser Ala Thr Cys Ser Ala Asp Glu Val Asp Glu Pro	
915 920 925	
gaa gga ggt gac aag ggc agg ctg ctg tca cca agg ccc agc ctc ctc	2832
Glu Gly Gly Asp Lys Gly Arg Leu Leu Ser Pro Arg Pro Ser Leu Leu	
930 935 940	
acc ccg gct gga gat ccc cgg gcc agt acc tcc ccc cag aag ccg ctg	2880
Thr Pro Ala Gly Asp Pro Arg Ala Ser Thr Ser Pro Gln Lys Pro Leu	
945 950 955 960	
gac ctg aag cag ctg aag cag cga gca gcc gcc atc ccc cct atc gtc	2928
Asp Leu Lys Gln Leu Lys Gln Arg Ala Ala Ala Ile Pro Pro Ile Val	
965 970 975	
acc aag gtc cat gag ccc ccc cgg gag gac aca gta ccc cca aag cca	2976
Thr Lys Val His Glu Pro Pro Arg Glu Asp Thr Val Pro Pro Lys Pro	
980 985 990	
gtt ccc cct gtg cct cca ccc acg cag cac cta cag cca gag ggt gac	3024
Val Pro Pro Val Pro Pro Pro Thr Gln His Leu Gln Pro Glu Gly Asp	
995 1000 1005	
gtg tct cag cag tcc gga gga agt cca cgt ggc aag tcc cgc agc cca	3072
Val Ser Gln Gln Ser Gly Gly Ser Pro Arg Gly Lys Ser Arg Ser Pro	
1010 1015 1020	
gtg cct cct gcc gag aaa gag gca gag aaa ccc gca ttc ttt ccg gct	3120
Val Pro Pro Ala Glu Lys Glu Ala Glu Lys Pro Ala Phe Phe Pro Ala	
1025 1030 1035 1040	
ttc cca act gag ggc caa agc tac cga ctg agc ccc cac gct ggt cat	3168
Phe Pro Thr Glu Gly Gln Ser Tyr Arg Leu Ser Pro His Ala Gly His	
1045 1050 1055	
cgg ctg cct tcc cat cct cca cgg gag gtg atc aag act tcc aca cgc	3216
Arg Leu Pro Ser His Pro Pro Arg Glu Val Ile Lys Thr Ser Thr Arg	
1060 1065 1070	
gct gac cct ctc ttc tcc tac aca ccc ccc ggt cac ccg ctg cct ctg	3264

CCP4 v1.3.0-64bit

Ala	Asp	Pro	Leu	Phe	Ser	Tyr	Thr	Pro	Pro	Gly	His	Pro	Leu	Pro	Leu	
1075						1080			1085							
ggc	ctc	cac	gat	agt	gcc	cgg	ccc	gtc	ctg	cca	cgt	ccc	ccc	atc	tct	3312
Gly	Leu	His	Asp	Ser	Ala	Arg	Pro	Val	Leu	Pro	Arg	Pro	Pro	Ile	Ser	
1090						1095			1100							
aac	ccc	cca	ccc	ctc	atc	tcc	tct	gcc	aag	cat	ccc	ggc	gta	ctt	gag	3360
Asn	Pro	Pro	Pro	Leu	Ile	Ser	Ser	Ala	Lys	His	Pro	Gly	Val	Leu	Glu	
1105						1110			1115			1120				
agg	cag	ctg	ggt	gcc	atc	tcc	cag	ggg	atg	tca	gtc	cag	ctt	cgt	gtg	3408
Arg	Gln	Leu	Gly	Ala	Ile	Ser	Gln	Gly	Met	Ser	Val	Gln	Leu	Arg	Val	
			1125						1130			1135				
cct	cac	tca	gag	cat	gcc	aag	ccc	atg	ggc	cct	ctc	acc	atg	gag	ctg	3456
Pro	His	Ser	Glu	His	Ala	Lys	Pro	Met	Gly	Pro	Leu	Thr	Met	Glu	Leu	
			1140						1145			1150				
ccc	ctt	gcc	gtg	gac	cct	aag	aag	ctg	ggg	aca	gca	ctg	gct	cgg	cca	3504
Pro	Leu	Ala	Val	Asp	Pro	Lys	Lys	Leu	Gly	Thr	Ala	Leu	Ala	Pro	Pro	
1155						1160						1165				
cca	gtg	gaa	gca	tca	cca	agg	gcc	tcc	cag	tac	ccg	ggc	tgc	aga	cgg	3552
Pro	Val	Glu	Ala	Ser	Pro	Arg	Ala	Ser	Gln	Tyr	Pro	Gly	Cys	Arg	Arg	
1170						1175						1180				
ccc	cag	cta	cag	agg	ctc	tat	cac	cca	cgc	acg	ccc	gca	gac	gtc	ctc	3600
Pro	Gln	Leu	Gln	Arg	Leu	Tyr	His	Pro	Arg	Thr	Pro	Ala	Asp	Val	Leu	
1185						1190						1195			1200	
tac	aag	ggt	acc	atc	agc	agg	atc	gtc	ggt	gag	gac	agc	cca	agt	cgc	3648
Tyr	Lys	Gly	Thr	Ile	Ser	Arg	Ile	Val	Gly	Glu	Asp	Ser	Pro	Ser	Arg	
			1205						1210						1215	
ctt	gac	cgg	gca	cga	gag	gac	acc	ctg	ccc	aag	ggc	cat	gtc	atc	tat	3696
Leu	Asp	Arg	Ala	Arg	Glu	Asp	Thr	Leu	Pro	Lys	Gly	His	Val	Ile	Tyr	
			1220						1225						1230	
gag	ggc	aag	aaa	ggc	cac	gtc	cta	tcc	tat	gaa	ggt	ggt	atg	tcc	gtg	3744
Glu	Gly	Lys	Lys	Gly	His	Val	Leu	Ser	Tyr	Glu	Gly	Gly	Met	Ser	Val	
1235						1240						1245				
tca	cag	tgc	tct	aag	gag	gat	gga	agg	agc	agc	tcg	ggc	cca	ccc	cat	3792
Ser	Gln	Cys	Ser	Lys	Glu	Asp	Gly	Arg	Ser	Ser	Ser	Gly	Pro	Pro	His	
1250						1255						1260				
gag	act	gcc	gcc	cct	aaa	cgc	acc	tat	gac	atg	atg	gag	ggc	cgt	gta	3840
Glu	Thr	Ala	Ala	Pro	Lys	Arg	Thr	Tyr	Asp	Met	Met	Glu	Gly	Arg	Val	
1265						1270						1275			1280	
ggc	agg	act	gtc	acc	tca	gcc	agc	ata	gag	gga	ctc	atg	ggc	cgc	gcc	3888
Gly	Arg	Thr	Val	Thr	Ser	Ala	Ser	Ile	Glu	Gly	Leu	Met	Gly	Arg	Ala	
			1285						1290						1295	
atc	cct	gag	cag	cac	agc	ccc	cac	ctc	aag	gag	cag	cat	cac	atc	cga	3936
Ile	Pro	Glu	Gln	His	Ser	Pro	His	Leu	Lys	Glu	Gln	His	His	Ile	Arg	

1315					1320					1325						
gac	tac	tta	cgg	cgg	gag	gcc	aag	ctc	ttg	aag	cga	gaa	ggg	aca	cca	4032
Asp	Tyr	Leu	Arg	Arg	Glu	Ala	Lys	Leu	Leu	Lys	Arg	Glu	Gly	Thr	Pro	
1330					1335					1340						
cct	ccc	cca	cca	cca	cct	cgg	gac	ctg	act	gag	acc	tac	aag	ccc	cgg	4080
Pro	Pro	Pro	Pro	Pro	Pro	Arg	Asp	Leu	Thr	Glu	Thr	Tyr	Lys	Pro	Arg	
1345					1350					1355					1360	
ccc	ctg	gac	cct	ctg	ggt	ccc	ctg	aag	ctg	aag	cgg	act	cac	gag	ggt	4128
Pro	Leu	Asp	Pro	Leu	Gly	Pro	Leu	Lys	Leu	Lys	Pro	Thr	His	Glu	Gly	
1365					1370					1375						
gtg	gta	gca	act	gtg	aag	gag	gcg	ggc	cgc	tct	atc	cat	gag	atc	cgg	4176
Val	Val	Ala	Thr	Val	Lys	Glu	Ala	Gly	Arg	Ser	Ile	His	Glu	Ile	Pro	
1380					1385					1390						
aga	gag	gag	ctg	cgc	cgc	aca	cct	gag	cta	ccc	ctg	gca	cca	cgg	cct	4224
Arg	Glu	Glu	Leu	Arg	Arg	Thr	Pro	Glu	Leu	Pro	Leu	Ala	Pro	Arg	Pro	
1395					1400					1405						
ctg	aag	gag	ggt	tcc	atc	acc	cag	ggc	acc	cca	ctc	aag	tac	gac	tct	4272
Leu	Lys	Glu	Gly	Ser	Ile	Thr	Gln	Gly	Thr	Pro	Leu	Lys	Tyr	Asp	Ser	
1410					1415					1420						
ggg	gca	ccc	tcc	act	ggc	acc	aag	aaa	cac	gac	gtg	cgc	tcc	atc	atc	4320
Gly	Ala	Pro	Ser	Thr	Gly	Thr	Lys	Lys	His	Asp	Val	Arg	Ser	Ile	Ile	
1425					1430					1435					1440	
ggc	agc	ccc	ggc	cgg	cct	ttc	cct	gcc	ctg	cac	cgg	ctg	gac	ata	atg	4368
Gly	Ser	Pro	Gly	Arg	Pro	Phe	Pro	Ala	Leu	His	Pro	Leu	Asp	Ile	Met	
1445					1450					1455						
gct	gac	gcc	cgg	gca	ctg	gag	cgt	gcc	tgc	tat	gaa	gag	agt	ctg	aag	4416
Ala	Asp	Ala	Arg	Ala	Leu	Glu	Arg	Ala	Cys	Tyr	Glu	Glu	Ser	Leu	Lys	
1460					1465					1470						
agc	cgg	tca	ggg	acc	agc	agt	ggt	gca	ggg	ggc	tcc	atc	aca	cgt	ggg	4464
Ser	Arg	Ser	Gly	Thr	Ser	Ser	Gly	Ala	Gly	Gly	Ser	Ile	Thr	Arg	Gly	
1475					1480					1485						
gct	cca	gtc	gtc	gtg	cct	gaa	ctg	ggc	aag	cca	cgg	caa	agc	cca	ctg	4512
Ala	Pro	Val	Val	Val	Pro	Glu	Leu	Gly	Lys	Pro	Arg	Gln	Ser	Pro	Leu	
1490					1495					1500						
act	tac	gaa	gac	cac	ggg	gca	ccc	ttc	acc	agt	cac	ctg	cca	cgt	ggc	4560
Thr	Tyr	Glu	Asp	His	Gly	Ala	Pro	Phe	Thr	Ser	His	Leu	Pro	Arg	Gly	
1505					1510					1515					1520	
tcc	cct	gtg	acc	acg	agg	gag	ccc	acg	cca	cgc	ctt	cag	gaa	ggc	agc	4608
Ser	Pro	Val	Thr	Thr	Arg	Glu	Pro	Thr	Pro	Arg	Leu	Gln	Glu	Gly	Ser	
1525					1530					1535						
ctc	cta	tcc	agc	aag	gcg	tcc	cag	gac	cgg	aag	ctg	aca	tct	aca	ccc	4656
Leu	Leu	Ser	Ser	Lys	Ala	Ser	Gln	Asp	Arg	Lys	Leu	Thr	Ser	Thr	Pro	
1540					1545					1550						
cgg	gag	atc	gcc	aag	tcc	cca	cac	agc								

cac ccc atc tcc ccc tat gag cac ttg ctc cgg ggc gtg act ggt gtg	4752
His Pro Ile Ser Pro Tyr Glu His Leu Leu Arg Gly Val Thr Gly Val	
1570 1575 1580	
gac ctg tac cgt ggt cac atc cca ttg gcc ttt gac ccc acc tcc ata	4800
Asp Leu Tyr Arg Gly His Ile Pro Leu Ala Phe Asp Pro Thr Ser Ile	
1585 1590 1595 1600	
ccc cga ggg atc cct ctg gaa gca gca gcc gca gcc tac tac ctg ccc	4848
Pro Arg Gly Ile Pro Leu Glu Ala Ala Ala Ala Tyr Tyr Leu Pro	
1605 1610 1615	
cgg cac ttg gcc ccc agc ccc acc tac cca cac ctg tac cca cct tac	4896
Arg His Leu Ala Pro Ser Pro Thr Tyr Pro His Leu Tyr Pro Pro Tyr	
1620 1625 1630	
ctc atc cgc ggc tac cct gac acg gcg gcc ctg gag aac cgc cag acc	4944
Leu Ile Arg Gly Tyr Pro Asp Thr Ala Ala Leu Glu Asn Arg Gln Thr	
1635 1640 1645	
atc atc aat gac tac atc acc tcg cag cag atg cac cac aac gct gcc	4992
Ile Ile Asn Asp Tyr Ile Thr Ser Gln Gln Met His His Asn Ala Ala	
1650 1655 1660	
tcc gcc atg gcc cag cgt gct gac atg ctg agg ggt ctg tca ccg cga	5040
Ser Ala Met Ala Gln Arg Ala Asp Met Leu Arg Gly Leu Ser Pro Arg	
1665 1670 1675 1680	
gag tcc tcg ctg gcc ctc aat tat tcc gct ggc cca aga ggc att atc	5088
Glu Ser Ser Leu Ala Leu Asn Tyr Ser Ala Gly Pro Arg Gly Ile Ile	
1685 1690 1695	
gac ctg tcc caa gtg cca cac ctg ccc gtg ctg gtg cca cca acg cca	5136
Asp Leu Ser Gln Val Pro His Leu Pro Val Leu Val Pro Pro Thr Pro	
1700 1705 1710	
ggc acc cct gcc acc gcc atc gac cgc ctt gcc tac ctc ccc act gcg	5184
Gly Thr Pro Ala Thr Ala Ile Asp Arg Leu Ala Tyr Leu Pro Thr Ala	
1715 1720 1725	
ccc cca ccc ttc agc agc cgc cac agt agc tca ccg ctg tcc cca gga	5232
Pro Pro Pro Phe Ser Ser Arg His Ser Ser Ser Pro Leu Ser Pro Gly	
1730 1735 1740	
ggc ccc act cac cta gct aaa cca act gcc aca tct tca tcg gag cgg	5280
Gly Pro Thr His Leu Ala Lys Pro Thr Ala Thr Ser Ser Ser Glu Arg	
1745 1750 1755 1760	
gaa cgg gaa cgt gag cgg gaa cga gac aag tcc atc ctc acg tct acc	5328
Glu Arg Glu Arg Glu Arg Glu Arg Asp Lys Ser Ile Leu Thr Ser Thr	
1765 1770 1775	
act aca gtg gag cat gca ccc atc tgg aga cct ggt acg gag cag agc	5376
Thr Thr Val Glu His Ala Pro Ile Trp Arg Pro Gly Thr Glu Gln Ser	
1780 1785 1790	
agc ggg gct ggg ggc agc agc cgc ccc gcc tcc cac acc cac cag cac	5424
Ser Gly Ala Gly Gly Ser Ser Arg Pro Ala Ser His Thr His Gln His	
1795 1800 1805	

000404-032704

tcg ccc atc tcc ccc cgg acc cag gac gcc ttg cag cag agg ccc agt	5472
Ser Pro Ile Ser Pro Arg Thr Gln Asp Ala Leu Gln Gln Arg Pro Ser	
1810 1815 1820	
gtg ctg cac aac acg agc atg aag ggc gtg gtc acc tcc gtg gaa ccc	5520
Val Leu His Asn Thr Ser Met Lys Gly Val Val Thr Ser Val Glu Pro	
1825 1830 1835 1840	
ggc acg ccc acg gtc ctg agg tcc acc tcc acc tct tcg cct gtc cgc	5568
Gly Thr Pro Thr Val Leu Arg Ser Thr Ser Thr Ser Ser Pro Val Arg	
1845 1850 1855	
cca gct gcc aca ttc cca cct gcc acc cac tgc cca ctt ggt ggc acc	5616
Pro Ala Ala Thr Phe Pro Pro Ala Thr His Cys Pro Leu Gly Gly Thr	
1860 1865 1870	
ctt gaa ggg gtc tac cct acc ctc atg gag ccc gtc ctg tta ccc aag	5664
Leu Glu Gly Val Tyr Pro Thr Leu Met Glu Pro Val Leu Leu Pro Lys	
1875 1880 1885	
gag acc tct cgg gtc gcc cgg ccc gag cgg ccc cgt gtg gac ggt ggc	5712
Glu Thr Ser Arg Val Ala Arg Pro Glu Arg Pro Arg Val Asp Gly Gly	
1890 1895 1900	
cat gcc ttc ctc acc aaa ccc ccg gcc cgg gag ccc gcc tcc tca ccc	5760
His Ala Phe Leu Thr Lys Pro Pro Ala Arg Glu Pro Ala Ser Ser Pro	
1905 1910 1915 1920	
agc aag agc tcc gag ccc cga tcc cta gca ccc ccc agc tcc agc cac	5808
Ser Lys Ser Ser Glu Pro Arg Ser Leu Ala Pro Pro Ser Ser Ser His	
1925 1930 1935	
aca gcc atc gcc cgc acc cca gca aag agc ctt gca ccc cac cat gcc	5856
Thr Ala Ile Ala Arg Thr Pro Ala Lys Ser Leu Ala Pro His His Ala	
1940 1945 1950	
agt ccg gac ccg ccg ggg ccc acc tcg gcc tca gat ctg cac cga gaa	5904
Ser Pro Asp Pro Pro Gly Pro Thr Ser Ala Ser Asp Leu His Arg Glu	
1955 1960 1965	
aag act caa agt aaa ccc ttt tcc atc cag gaa ttg gaa ctc cgt tct	5952
Lys Thr Gln Ser Lys Pro Phe Ser Ile Gln Glu Leu Glu Leu Arg Ser	
1970 1975 1980	
ctg ggt tac cac agt gga gct ggc tac agc ccc gat ggg gtg gag ccc	6000
Leu Gly Tyr His Ser Gly Ala Gly Tyr Ser Pro Asp Gly Val Glu Pro	
1985 1990 1995 2000	
atc agc ccg gtg agc tcc ccc agc ctg acc cac gac aag ggg ctc tcc	6048
Ile Ser Pro Val Ser Ser Pro Ser Leu Thr His Asp Lys Gly Leu Ser	
2005 2010 2015	
aaa cct ctg gaa gag cta gag aag agc cac ttg gaa ggg gag ctg cgg	6096
Lys Pro Leu Glu Glu Leu Glu Lys Ser His Leu Glu Gly Glu Leu Arg	
2020 2025 2030	
cac aag cag cca ggc ccc atg aag ctc agc gcg gag gct gcc cat ctc	6144
His Lys Gln Pro Gly Pro Met Lys Leu Ser Ala Glu Ala Ala His Leu	
2035 2040 2045	
cca cat ctg cgg cca ctg ccc gag agc cag ccc tca tcc agc cca ctc	6192

10220-1070600

Pro	His	Leu	Arg	Pro	Leu	Glu	Ser	Gln	Pro	Ser	Ser	Pro	Leu				
2050						2055			2060								
ctc	cag	act	gcc	cca	ggc	atc	aaa	ggg	cac	cag	agg	gtg	gtc	acc	ctg		6240
Leu	Gln	Thr	Ala	Pro	Gly	Ile	Lys	Gly	His	Gln	Arg	Val	Val	Thr	Leu		
2065					2070				2075						2080		
gct	cag	cac	atc	agc	gag	gtc	att	acg	cag	gac	tac	acc	cgg	cac	cac		6288
Ala	Gln	His			Ser	Glu	Val	Ile	Thr	Gln	Asp	Tyr	Thr	Arg	His	His	
				2085					2090						2095		
ccg	cag	cag	ctc	agt	ggc	ccc	ctt	ccc	gcc	cct	ctc	tac	tcc	ttt	ccc		6336
Pro	Gln	Gln	Leu	Ser	Gly	Pro	Leu	Pro	Ala	Pro	Leu	Tyr	Ser	Phe	Pro		
			2100					2105					2110				
gga	gcc	agc	tgc	cct	gtg	ctg	gat	ctt	cgc	cgc	cca	ccc	agt	gac	ctc		6384
Gly	Ala	Ser	Cys	Pro	Val	Leu	Asp	Leu	Arg	Arg	Pro	Pro	Ser	Asp	Leu		
		2115				2120						2125					
tac	ctc	cca	ccc	ccc	gac	cat	ggc	acc	cca	gcc	cgg	gga	tcc	ccc	cac		6432
Tyr	Leu	Pro	Pro	Pro	Asp	His	Gly	Thr	Pro	Ala	Arg	Gly	Ser	Pro	His		
	2130					2135					2140						
agt	gaa	ggg	ggc	aaa	agg	tcc	cca	gaa	ccc	agc	aaa	aca	tcg	gtc	ctg		6480
Ser	Glu	Gly	Gly	Lys	Arg	Ser	Pro	Glu	Pro	Ser	Lys	Thr	Ser	Val	Leu		
2145					2150					2155					2160		
ggc	agc	agt	gag	gat	gcc	att	gag	cct	gtg	tcc	cca	cca	gag	ggc	atg		6528
Gly	Ser	Ser	Glu	Asp	Ala	Ile	Glu	Pro	Val	Ser	Pro	Pro	Glu	Gly	Met		
			2165						2170					2175			
act	gag	cca	gga	cat	gct	cgg	agc	gct	gtg	tac	cca	ctg	ctg	tat	cga		6576
Thr	Glu	Pro	Gly	His	Ala	Arg	Ser	Ala	Val	Tyr	Pro	Leu	Leu	Tyr	Arg		
		2180						2185				2190					
gac	ggg	gaa	cag	ggc	gag	ccc	agg	atg	ggc	tct	aag	tct	cca	ggc	aac		6624
Asp	Gly	Glu	Gln	Gly	Glu	Pro	Arg	Met	Gly	Ser	Lys	Ser	Pro	Gly	Asn		
		2195					2200				2205						
acc	agc	cag	ccg	cca	gcc	ttc	ttc	agt	aag	ctg	act	gag	agc	aac	tcc		6672
Thr	Ser	Gln	Pro	Pro	Ala	Phe	Phe	Ser	Lys	Leu	Thr	Glu	Ser	Asn	Ser		
	2210					2215					2220						
gcc	atg	gtg	aag	tcg	aag	aag	cag	gag	atc	aac	aag	aaa	ctc	aac	acc		6720
Ala	Met	Val	Lys	Ser	Lys	Lys	Gln	Glu	Ile	Asn	Lys	Lys	Leu	Asn	Thr		
2225					2230				2235						2240		
cac	aac	cgg	aac	gag	cca	gaa	tac	aat	att	ggc	cag	cct	ggg	acg	gaa		6768
His	Asn	Arg	Asn	Glu	Pro	Glu	Tyr	Asn	Ile	Gly	Gln						

ccg	ccg	ctc	ggc	gcc	aat	gct	ttt	aac	cct	ctg	aat	gcc	agc	gcc	agt	6960
Pro	Pro	Leu	Gly	Ala	Asn	Ala	Phe	Asn	Pro	Leu	Asn	Ala	Ser	Ala	Ser	
2305					2310					2315					2320	
ctg	ccc	gct	gct	gct	atg	ccc	ata	acc	act	gct	gac	gga	cgg	agt	gac	7008
Leu	Pro	Ala	Ala	Ala	Met	Pro	Ile	Thr	Thr	Ala	Asp	Gly	Arg	Ser	Asp	
				2325					2330						2335	
cac	gca	ctc	acc	tcg	cca	ggg	gga	ggg	ggg	aaa	gcc	aag	gtc	tct	ggc	7056
His	Ala	Leu	Thr	Ser	Pro	Gly	Gly	Gly	Gly	Lys	Ala	Lys	Val	Ser	Gly	
			2340					2345					2350			
aga	cct	agc	agc	cga	aaa	gcc	aag	tcg	cca	gca	cca	ggc	cta	gcg	tcc	7104
Arg	Pro	Ser	Ser	Arg	Lys	Ala	Lys	Ser	Pro	Ala	Pro	Gly	Leu	Ala	Ser	
		2355					2360					2365				
gga	gac	cga	ccc	cct	tct	gtc	tcc	tca	gta	cac	tca	gag	ggg	gac	tgc	7152
Gly	Asp	Arg	Pro	Pro	Ser	Val	Ser	Ser	Val	His	Ser	Glu	Gly	Asp	Cys	
	2370					2375					2380					
aat	cgc	cga	aca	cca	ctc	acc	aac	cgt	gtg	tgg	gag	gac	cgg	ccc	tca	7200
Asn	Arg	Arg	Thr	Pro	Leu	Thr	Asn	Arg	Val	Trp	Glu	Asp	Arg	Pro	Ser	
2385					2390					2395					2400	
tct	gca	ggg	tcc	acg	cca	ttc	ccc	tac	aac	cct	ttg	att	atg	agg	cta	7248
Ser	Ala	Gly	Ser	Thr	Pro	Phe	Pro	Tyr	Asn	Pro	Leu	Ile	Met	Arg	Leu	
			2405						2410						2415	
cag	gca	ggg	gtc	atg	gcc	tcc	ccg	ccc	cca	cct	ggc	ctt	gcg	gca	ggc	7296
Gln	Ala	Gly	Val	Met	Ala	Ser	Pro	Pro	Pro	Pro	Gly	Leu	Ala	Ala	Gly	
		2420						2425					2430			
agc	ggg	ccc	cta	gct	ggg	ccc	cac	cac	gcc	tgg	gat	gag	gag	ccc	aag	7344
Ser	Gly	Pro	Leu	Ala	Gly	Pro	His	His	Ala	Trp	Asp	Glu	Glu	Pro	Lys	
		2435					2440					2445				
cca	ctg	ctg	tgt	tca	cag	tat	gag	aca	ctc	tcg	gac	agc	gag			7386
Pro	Leu	Leu	Cys	Ser	Gln	Tyr	Glu	Thr	Leu	Ser	Asp	Ser	Glu			
	2450					2455					2460					